

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2004, 23:34:58 ; Search time 158 Seconds

(without alignments)
615.289 Million cell updates/sec

Title: US-10-090-185-9

Perfect score: 1388

Sequence: 1 RCLMEBSRLQTAATAAAGG.....LNTQKIKVCIDKSDGVAA 271

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: geneeqp19808:*\n2: geneeqp19908:*\n3: geneeqp20008:*\n4: geneeqp20018:*\n5: geneeqp20028:*\n6: geneeqp20038:*\n7: geneeqp20038:*\n8: geneeqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1388	100.0	271	4	AAV72841	AAV72841 Mouse Sta
2	1388	100.0	770	2	AAV72082	AAV72082 Mouse Sta
3	1388	100.0	770	2	AAW03176	AAW03176 Mouse Sta
4	1377	99.2	720	5	AAE22055	AAE22055 Human Sta
5	1377	99.2	769	5	ABH57164	ABH57164 Mouse isc
6	1377	99.2	769	5	AAE22054	AAE22054 Human Sta
7	1377	99.2	769	5	AAE22056	AAE22056 Human pro
8	1377	99.2	770	2	AAE22995	AAE22995 Mouse liv
9	1377	99.2	770	2	AAV03768	AAV03768 Human STA
10	1377	99.2	770	2	AAH12377	AAH12377 N-termina
11	1377	99.2	770	2	AAE14652	AAE14652 Murine ST
12	1377	99.2	770	5	ABG69497	ABG69497 Human bal
13	1377	99.2	770	5	ABU10476	ABU10476 Mouse STA
14	1377	99.2	770	8	ADN04365	ADN04365 Antisori
15	1377	99.2	770	8	AAE58442	AAE58442 Lung canc
16	1377	99.0	770	7	ADD44738	ADD44738 Rat Prote
17	1372	98.8	770	2	AAE82993	AAE82993 Human pla
18	1372	98.8	770	2	AAH19964	AAH19964 Human s1g
19	1372	98.8	770	5	AAE15174	AAE15174 Human Sta
20	1372	98.8	770	7	ADD44740	ADD44740 Human pro
21	1290	92.9	252	4	AAV72846	AAV72846 Mouse Sta
22	1212	87.3	236	4	AAV72847	AAV72847 Mouse Sta
23	1172	84.4	229	4	AAV72850	AAV72850 Mouse Sta
24	1168	84.1	229	4	AAV72863	AAV72863 Mouse Sta
25	1167	84.1	229	4	AAV72862	AAV72862 Mouse Sta

26	1157.5	83.4	228	4	AAV72861	AAV72861 Mouse Sta
27	1146	82.6	229	4	AAV72860	AAV72860 Mouse Sta
28	1142	82.3	223	4	AAV72854	AAV72854 Mouse Sta
29	1094	78.8	213	4	AAV72851	AAV72851 Mouse Sta
30	943	67.9	185	4	AAV72855	AAV72855 Mouse Sta
31	899	64.8	176	4	AAV72848	AAV72848 Mouse Sta
32	723	52.1	143	4	AAV72849	AAV72849 Mouse Sta
33	669	48.2	129	4	AAV72856	AAV72856 Mouse Sta
34	654	47.1	749	5	AAV72852	AAV72852 Mouse Sta
35	653	47.0	128	4	AAV72852	AAV72852 Mouse Sta
36	647	46.6	268	4	AAV72844	AAV72844 Mouse Sta
37	647	46.6	680	6	ABR59713	ABR59713 Human s1g
38	647	46.6	712	2	AAV72079	AAV72079 Human Sta
39	647	46.6	712	2	AAW03170	AAW03170 Human STA
40	647	46.6	712	2	AAW62995	AAW62995 Human Sta
41	647	46.6	712	6	ABU04747	ABU04747 Human exp
42	647	46.6	712	6	ABU04735	ABU04735 Human exp
43	647	46.6	712	6	ABU04745	ABU04745 Human exp
44	647	46.6	712	6	ABU04743	ABU04743 Human exp
45	647	46.6	712	8	ADH57035	ADH57035 Truncated

ALIGNMENTS

RESULT 1	
AAV72841	
ID	AAV72841 standard; protein; 271 AA.
XX	
AC	AAV72841;
XX	
DT	31-MAY-2001 (first entry)
XX	
DE	Mouse Stat3 protein fragment #2 (107-377 amino acids).
XX	
KW	Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription; cellular transformation; dysproliferative disease; cancer; psoriasis; therapy.
KW	
XX	
OS	Mus musculus.
XX	
FN	Key
FT	Region
FT	24..48
FT	/note="Stat3-C-Jun interaction region 1; corresponds to 130-154 position of Stat3 protein"
FT	236..252
FT	/note="Stat3-C-Jun interaction region 2; corresponds to 342-358 position of Stat3 protein"
XX	
XX	WO200116605-A2.
XX	
XX	08-MAR-2001.
XX	
XX	30-AUG-2000; 2000WO-US023822.
XX	
XX	31-AUG-1999; 99US-00387418.
XX	
XX	(UYRQ) UNIV ROCKEFELLER.
XX	
XX	Zhang X, Horvath C, Wzreszcynska MH, Darnell JB;
XX	
XX	WPI; 2001-226705/23.
XX	
XX	Identifying an agent for use in modulating the interaction between
XX	transcription factor c-Jun and a Stat3 protein.
XX	
XX	Claim 65; Page 67-68; 86pp; English.
XX	
XX	The present sequence is mouse Stat3 protein fragment containing 107-377
XX	amino acids of Stat3 protein. This Stat3 fragment showed strong binding
XX	to c-Jun protein in the cell extract. The invention relates to methods
XX	for identifying interacting regions of transcription factors and methods
XX	for identifying agents which modulates the interaction between a

transcription factor such as c-Jun and a Stat protein such as Stat-1 and Stat-3, useful for modulating gene transcription e.g., cellular transformation. These identifying agents are used in the treatment of CC dysproliferative diseases and also for treating cancer and psoriasis. A CC Stat protein comprises the N-terminal domain, coiled-coil domain, DNA binding domain, linker domain, SH2 domain and transactivation domain CC XX

Sequence 271 AA;
Query Match 100.0%; Score 1388; DB 4; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.5e-116;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RCIMESRLIQTATATAAQQGGQANHPAAVTEKQMLEQHLQDVRKRVDDLEQKMKVE 60
DB 1 RCIMESRLIQTATATAAQQGGQANHPAAVTEKQMLEQHLQDVRKRVDDLEQKMKVE 60
QY 61 NLDDDPENFKTKSGQDMODLNGNNSVTRQKQOLEQMLTALDQRRSIVSELAGLLS 120
DB 61 NLDDDPENFKTKSGQDMODLNGNNSVTRQKQOLEQMLTALDQRRSIVSELAGLLS 120
QY 121 AMEYVOKTLTDEELADMKRRPEIACIGPPNICIDRLNNWITSLAESQLQTRQOIKKLE 180
DB 121 AMEYVOKTLTDEELADMKRRPEIACIGPPNICIDRLNNWITSLAESQLQTRQOIKKLE 180
QY 181 LOOKSVYKGPPIYQHRPMLERIVELFRNLKMSAFVVERQPCMPHMDRPLVKTGVQFT 240
DB 181 LOOKSVYKGPPIYQHRPMLERIVELFRNLKMSAFVVERQPCMPHMDRPLVKTGVQFT 240
QY 241 TKVRLVKKPELNYQIKIVCICDKSGDVAA 271
DB 241 TKVRLVKKPELNYQIKIVCICDKSGDVAA 271

RESULT 2

AAR72082
ID AAR72082 standard; protein; 770 AA.

XX AAR72082;

DT 25-MAR-2003 (revised)
DT 27-SEP-1995 (first entry)

XX Mouse Stat3 (1996f).

XX Signal transducer and activator of transcription; STAT; 1996f; Stat3;
KW receptor recognition factor; transcription factor; cellular debilitation;
KW derangement; dysfunction; interferon-gamma.

OS Mus sp.

XX WO9508629-A1.

XX 30-MAR-1995.

PF 26-SEP-1994; 94WO-US010849.

PR 24-SEP-1993; 93US-00126588.

PR 24-SEP-1993; 93US-00126595.

PR 11-MAR-1994; 94US-00212184.

PR 11-MAR-1994; 94US-00212185.

XX (UYRQ) UNIV ROCKEFELLER.

XX Darnell JE, Schindler CW, Shuai K, Wen Z, Zhong Z;

XX WPI; 1995-139598/18.

XX N-PSDB; AAQ89340.

XX Receptor recognition factor implicated in transcriptional stimulation of
PT genes - useful in drug screening assays and/or for treating cellular
PT debilitations, derangements and/or dysfunctions, etc.

PS Claim 1; Page 107-110; 160pp; English.

XX A fragment encoding the human Stat3 protein was used to screen a murine
CC thymus and spleen cDNA for homologous proteins. A highly homologous gene
CC (given in AAQ89338) was isolated that encoded a 91 kDa protein (AAR72080)
CC (Stat3) that was responsive to interferon-gamma. Using a fragment of the
CC mouse gene as probe, 2 additional members of the 113-91 family of
CC receptor recognition factor proteins were isolated. The 2 genes (AAQ89339
CC -40) were cloned in plasmids 136f1 and 198f6 and encoded proteins termed
CC Stat4 (AAR72081) and Stat3 (AAR72082), respectively. (Updated on 25-MAR-
CC 2003 to correct PN field.)

XX Sequence 770 AA;

Query Match 100.0%; Score 1388; DB 2; Length 770;
Best Local Similarity 100.0%; Pred. No. 1e-115;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RCIMESRLIQTATATAAQQGGQANHPAAVTEKQMLEQHLQDVRKRVDDLEQKMKVE 60
DB 107 RCIMESRLIQTATATAAQQGGQANHPAAVTEKQMLEQHLQDVRKRVDDLEQKMKVE 166
QY 61 NLDDDPENFKTKSGQDMODLNGNNSVTRQKQOLEQMLTALDQRRSIVSELAGLLS 120
DB 61 NLDDDPENFKTKSGQDMODLNGNNSVTRQKQOLEQMLTALDQRRSIVSELAGLLS 120
QY 167 NLDDDPENFKTKSGQDMODLNGNNSVTRQKQOLEQMLTALDQRRSIVSELAGLLS 226
DB 167 NLDDDPENFKTKSGQDMODLNGNNSVTRQKQOLEQMLTALDQRRSIVSELAGLLS 226
QY 121 AMEYVOKTLTDEELADMKRRPEIACIGPPNICIDRLNNWITSLAESQLQTRQOIKKLE 180
DB 121 AMEYVOKTLTDEELADMKRRPEIACIGPPNICIDRLNNWITSLAESQLQTRQOIKKLE 180
QY 227 AMEYVOKTLTDEELADMKRRPEIACIGPPNICIDRLNNWITSLAESQLQTRQOIKKLE 286
DB 227 AMEYVOKTLTDEELADMKRRPEIACIGPPNICIDRLNNWITSLAESQLQTRQOIKKLE 286
QY 181 LOOKSVYKGPPIYQHRPMLERIVELFRNLKMSAFVVERQPCMPHMDRPLVKTGVQFT 240
DB 181 LOOKSVYKGPPIYQHRPMLERIVELFRNLKMSAFVVERQPCMPHMDRPLVKTGVQFT 240
QY 287 LOOKSVYKGPPIYQHRPMLERIVELFRNLKMSAFVVERQPCMPHMDRPLVKTGVQFT 346
DB 287 LOOKSVYKGPPIYQHRPMLERIVELFRNLKMSAFVVERQPCMPHMDRPLVKTGVQFT 346
QY 241 TKVRLVKKPELNYQIKIVCICDKSGDVAA 271
DB 241 TKVRLVKKPELNYQIKIVCICDKSGDVAA 271

RESULT 3

AAW03176
ID AAW03176 standard; protein; 770 AA.

XX AAW03176;

DT 24-OCT-1996 (first entry)

XX Mouse STAT4.

XX STAT; STAT4; signal transducer and activator of transcription;
KW DNA binding protein; ligand; receptor; oncogenesis; inflammation;
KW autoimmune disease; antagonist; therapy.

OS Mus sp.

XX Key Location/Qualifiers

FT Domain 398..508

FT /label= DNA binding domain

XX /note= "Claim 3, page 110"

XX WO9620954-A2.

XX 11-JUL-1996.

PF 28-DEC-1995; 95WO-US017025.

PR 06-JAN-1995; 95US-0036796.

XX (UYRQ) UNIV ROCKEFELLER.

XX Darnell JE, Wen Z, Horvath CM, Zhong Z;

XX WPI; 1996-333941/33.

DR N-PSDB; AAT31280.
 XX New STAT protein DNA-binding domain peptide(s) - useful for diagnosing,
 PT preventing or treating cellular dysfunction, e.g. oncogenesis,
 PT inflammation, parasitic disease or autoimmunity.
 XX
 XX Disclosure; Page 87-90; 138pp; English.
 XX
 CC Mouse signal transducer and activator of transcription (STAT) protein
 CC STAT4 (AA031276) serves a dual purpose, i.e. signal transduction from
 CC ligand-activated receptor kinase complexes followed by nuclear
 CC translocation and DNA binding to activate transcription. Recombinant
 CC STAT4 can be obtained using cDNA clone 198f6 (AAT31278) obtained from
 CC splenic/thymic cells. STAT4 includes a DNA-binding domain (see also
 CC AA03167) capable of both receptor recognition and message delivery via
 CC DNA binding in a receptor-ligand specific manner. STAT proteins and their
 CC DNA binding domains (see also AA03165-75) are useful for screening
 CC antagonists used to inhibit STAT-mediated signal transduction and
 CC activation of transcription
 CC
 SQ Sequence 770 AA;
 Query Match 100.0%; Score 1388; DB 2; Length 770;
 Best Local Similarity 100.0%; Pred. No. 1e-115;
 Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RCLMESRLTQTAATTAAGGAGNPTAAVTEKQOMLEOHLDVREKVDLEQKRYVE 60
 DB 107 RCLMESRLTQTAATTAAGGAGNPTAAVTEKQOMLEOHLDVREKVDLEQKRYVE 166
 QY 61 NLQDDFDNRYKTKLSQGMODLNGNSVTRQKMOLEOMLTALDQMRSTVSELAGLLS 120
 DB 167 NLQDDFDNRYKTKLSQGMODLNGNSVTRQKMOLEOMLTALDQMRSTVSELAGLLS 226
 QY 121 AMEYVQKTLTBEALADMKRRRPRPACIGSPNLCIDRLRNWITSLSAESQLQTRQIKKLE 180
 DB 227 AMEYVQKTLTBEALADMKRRRPRPACIGSPNLCIDRLRNWITSLSAESQLQTRQIKKLE 286
 QY 181 LQOKSYKGDPIVQHRPMLSEIYELFNNLMSKAFVVEROPCMHPRPPLVIKYGVOFT 240
 DB 287 LQOKSYKGDPIVQHRPMLSEIYELFNNLMSKAFVVEROPCMHPRPPLVIKYGVOFT 346
 QY 241 TKVRLVKEPELNYQKIKVCIDKSGDVAA 271
 DB 347 TKVRLVKEPELNYQKIKVCIDKSGDVAA 377
 RESULT 4
 AAE22055 standard; protein; 720 AA.
 ID AAE22055
 AC AAE22055;
 XX
 XX 25-JUL-2002 (first entry)
 DT
 XX
 XX Human Stat3beta protein.
 DE
 XX
 XX Human; signal transducer and activator of transcription 3; ischaemia;
 KW immune response; Stat3; coronary atherosclerosis; vascular occlusion;
 KW hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
 KW inflammation; chronic obstructive pulmonary disease; cardiac arrest;
 KW insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
 KW shock; chronic active hepatitis; adult respiratory distress syndrome;
 KW nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
 KW Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
 KW polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
 KW proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
 KW carcinoma; degenerative disorder; gene therapy; growth deficiency;
 KW cirrhosis; hypoproliferative disorder; lesion; statbeta.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers

FT Misc-difference 713. 714
 FT /note= "Encoded by ACA CCA TTC"
 XX
 XX WO200220032-A1.
 PN
 XX
 XX 14-MAR-2002.
 PD
 XX
 XX 10-SEP-2001; 2001WO-US028254.
 PF
 XX
 XX 08-SEP-2000; 2000US-0231212P.
 PR
 XX
 XX (UYJO) UNIV JOHNS HOPKINS.
 PA (UYSF-) UNIV SOUTH FLORIDA.
 XX
 XX Yu H, Pardoll D, Jove R, Dalton W;
 PI
 XX
 XX WPI; 2002-362218/39.
 DR N-PSDB; AAD35066.
 XX
 PT Modulating angiogenesis and an immune response in an individual, for
 PT treating a hypoxic or ischemic condition, comprises administering a
 PT compound that modulates the activity of a signal transducer and activator
 PT of transcription 3.
 XX
 PS Disclosure; Page 87-89; 94pp; English.
 XX
 XX The invention relates to a method of modulating angiogenesis and immune
 CC response. Method involves administering to an individual a compound that
 CC modulate the activity of signal transducer and activator of transcription
 CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
 CC hypoxic or ischemic condition or disorder which is the result of stroke,
 CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
 CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
 CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
 CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
 CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
 CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
 CC with neovascularisation. Suppressing an immune response is useful for
 CC ameliorating a symptom of an autoimmune disease such as systemic lupus
 CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
 CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
 CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
 CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
 CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
 CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
 CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
 CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
 CC dense deposit disease. The method is useful in preventing or treating
 CC specific proliferative and oncogenic disease which includes sarcomas and
 CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
 CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
 CC hypoproliferative disorders, physical trauma, lesions and wounds. The
 CC method is also used in gene therapy. The present sequence is human
 CC Stat3beta protein
 CC
 SQ Sequence 720 AA;
 Query Match 99.2%; Score 1377; DB 5; Length 720;
 Best Local Similarity 99.3%; Pred. No. 9.1e-115;
 Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RCLMESRLTQTAATTAAGGAGNPTAAVTEKQOMLEOHLDVREKVDLEQKRYVE 60
 DB 107 RCLMESRLTQTAATTAAGGAGNPTAAVTEKQOMLEOHLDVREKVDLEQKRYVE 166
 QY 61 NLQDDFDNRYKTKLSQGMODLNGNSVTRQKMOLEOMLTALDQMRSTVSELAGLLS 120
 DB 167 NLQDDFDNRYKTKLSQGMODLNGNSVTRQKMOLEOMLTALDQMRSTVSELAGLLS 226
 QY 121 AMEYVQKTLTBEALADMKRRRPRPACIGSPNLCIDRLRNWITSLSAESQLQTRQIKKLE 180
 DB 227 AMEYVQKTLTBEALADMKRRRPRPACIGSPNLCIDRLRNWITSLSAESQLQTRQIKKLE 286

QY 181 LQOKVSYKDDPIVQHRPMLERIVELFRNLKMSAFVERQPCMPHDPRLVTKTGVQFT 240
 DB 287 LQOKVSYKDDPIVQHRPMLERIVELFRNLKMSAFVERQPCMPHDPRLVTKTGVQFT 346
 QY 241 TKVRLVKEPELNYOLKIVCICDKSGDVAA 271
 DB 347 TKVRLVKEPELNYOLKIVCICDKSGDVAA 377

RESULT 5

ABB57164
 ID ABB57164 standard; protein; 769 AA.

AC ABB57164;
 XX

DT 07-MAR-2002 (first entry)
 XX

DE Mouse ischaemic condition related protein sequence SEQ ID NO:398.
 XX

KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 KM vasospastic ischaemia; ischaemic condition; ischaemic disease.
 XX

OS Mus musculus.
 XX

PN WO200188188-A2.
 XX

PD 22-NOV-2001.
 XX

PF 18-MAY-2001; 2001WO-JP004192.
 XX

PR 18-MAY-2000; 2000JP-00145977.
 XX

PA (UYNI-) UNIV NITHON SCHOOL JURIDICAL PERSON.
 XX

PI Ishikawa K, Arai S, Takahashi Y, Nagata T, Ishii Y;
 XX

DR WPI, 2002-034733/04.
 XX

DR N-PSDB; ABI99454.
 XX

PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or by
 PT determining the expression profile of a gene group comprising these
 PT genes.

PS Claim 2; Page 1084-1087; 2690pp; English.
 XX

CC The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC gene (1) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (1). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
 CC expression levels of particular genes (ABI99202 to ABI99912, encoding the
 CC protein sequences in ABB57020 to ABB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The expression
 CC levels or expression profiles produced by these genes are used as an
 CC indicator when screening for ischaemic condition-improving drugs or
 CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
 CC primers for a mouse ischaemic condition related sequence, which are used
 CC in the exemplification of the present invention
 CC

XX
 SQ Sequence 769 AA;

Query Match 99.2%; Score 1377; DB 5; Length 769;
 Best Local Similarity 99.3%; Pred. No. 1e-114;

Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RCLWESRLQTRATTAAGCGGQANHTAAVTEKQMLRQHLQDVYKRVQDLEQKKVYE 60
 DB 107 RCLWESRLQTRATTAAGCGGQANHTAAVTEKQMLRQHLQDVYKRVQDLEQKKVYE 166
 QY 61 NIQDDPFFNYKTLKSGQDMQDINGNNSYTRQKMOQLRQMLTALDQMRSSIVSELAGILS 120

DB 167 NIQDDPFFNYKTLKSGQDMQDINGNNSYTRQKMOQLRQMLTALDQMRSSIVSELAGILS 226
 QY 121 AMEYVOKTLDLDELAAMKRRPEIACIGPPNICLDRLENWITSIAESQLOTRQOIKLEE 180
 DB 227 AMEYVOKTLDLDELAAMKRRQOACIGPPNICLDRLENWITSIAESQLOTRQOIKLEE 286
 QY 181 LQOKVSYKDDPIVQHRPMLERIVELFRNLKMSAFVERQPCMPHDPRLVTKTGVQFT 240
 DB 287 LQOKVSYKDDPIVQHRPMLERIVELFRNLKMSAFVERQPCMPHDPRLVTKTGVQFT 346
 QY 241 TKVRLVKEPELNYOLKIVCICDKSGDVAA 271
 DB 347 TKVRLVKEPELNYOLKIVCICDKSGDVAA 377

RESULT 6

AAE22054
 ID AAE22054 standard; protein; 769 AA.

AC AAE22054;
 XX

DT 25-JUL-2002 (first entry)
 XX

DE Human Stat3 protein.
 XX

KW Human; signal transducer and activator of transcription 3; ischaemia;
 KM immune response; Stat3; coronary atherosclerosis; vascular occlusion;
 KM hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
 KM inflammation; chronic obstructive pulmonary disease; cardiac arrest;
 KM insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
 KM shock; chronic active hepatitis; adult respiratory distress syndrome;
 KM nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
 KM Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
 KM polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
 KM proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
 KM carcinoma; degenerative disorder; gene therapy; growth deficiency;
 KM cirrhosis; hypoproliferative disorder; lesion.
 XX

OS Homo sapiens.
 XX

PN WO200220032-A1.
 XX

PD 14-MAR-2002.
 XX

PF 10-SEP-2001; 2001WO-US028254.
 XX

PR 08-SEP-2000; 2000US-0231212P.
 XX

PA (UYNO) UNIV JOHNS HOPKINS.
 XX

PI (UYSF-) UNIV SOUTH FLORIDA.
 XX

PI Yu H, Pardoll D, Jove R, Dalton W;
 XX

DR WPI; 2002-362218/39.
 XX

DR N-PSDB; AAD35065.
 XX

PT Modulating angiogenesis and an immune response in an individual, for
 PT treating a hypoxic or ischemic condition, comprises administering a
 PT compound that modulates the activity of a signal transducer and activator
 PT of transcription 3.
 PT

PS Disclosure; Page 83-85; 94pp; English.
 XX

CC The invention relates to a method of modulating angiogenesis and immune
 CC response. Method involves administering to an individual a compound that
 CC modulate the activity of signal transducer and activator of transcription
 CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
 CC hypoxic or ischaemic condition or disorder which is the result of stroke,
 CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
 CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
 CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
 CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
 CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest.

nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy with neovascularisation. Suppressing an immune response is useful for ameliorating a symptom of an autoimmune disease such as systemic lupus erythematosus, multiple sclerosis, insulin dependent diabetes mellitus, Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis, mixed connective tissue disease, primary biliary cirrhosis, pernicious anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo, gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis, idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and dense deposit disease. The method is useful in preventing or treating specific proliferative and oncogenic disease which includes sarcomas and carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia, fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency, hypoproliferative disorders, physical trauma, lesions and wounds. The method is also used in gene therapy. The present sequence is human Stat3 protein

Sequence 769 AA;

Query Match 99.2%; Score 1377; DB 5; Length 769;
Best Local Similarity 99.3%; Pred. No. 1e-114; Indels 0; Gaps 0;
Matches 269; Conservative 1; Mismatches 1;

1 RCLMEESRLQTAATAAQQGQANPTAAVTEKQOMLEOHLQVRRKRVQDLEQKKVVE 60
107 RCLMEESRLQTAATAAQQGQANPTAAVTEKQOMLEOHLQVRRKRVQDLEQKKVVE 166
QY 61 NIQDDPFRNYKTKSQGMODINGNNOSTYRQKQOLEBMTALDQMRSSYSELAGLIS 120
DB 167 NIQDDPFRNYKTKSQGMODINGNNOSTYRQKQOLEBMTALDQMRSSYSELAGLIS 226
QY 121 AMEYVQKTLTDEBLADWKRPRPACIGGPPNCLDRLEWITSLAESQQTQOQIKLEE 180
DB 227 AMEYVQKTLTDEBLADWKRPRPACIGGPPNCLDRLEWITSLAESQQTQOQIKLEE 286
QY 181 LQOKSYSGDPIVQHRPMLERIVELFRNLKMSAFVEROPCPMPHPRPRLVYKTVGQPT 240
DB 287 LQOKSYSGDPIVQHRPMLERIVELFRNLKMSAFVEROPCPMPHPRPRLVYKTVGQPT 346
QY 241 TKVRLVFPPELNYOLKIKVCIDKDSGVAA 271
DB 347 TKVRLVFPPELNYOLKIKVCIDKDSGVAA 377

RESULT 7

AAE22056
ID AAE22056 standard; protein; 769 AA.

AAE22056;

25-JUL-2002 (first entry)

Human protein related to angiogenesis regulation.

Human; signal transducer and activator of transcription 3; leukaemia; immune response; Stat3; coronary atherosclerosis; vascular occlusion; hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia; inflammation; chronic obstructive pulmonary disease; cardiac arrest; insulin dependent diabetes mellitus; emphysema; trauma; scleroderma; shock; chronic active hepatitis; adult respiratory distress syndrome; nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis; Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy; polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia; proliferative disease; Grave's disease; ulcerative colitis; sarcoma; carcinoma; degenerative disorder; gene therapy; growth deficiency; cirrhosis; hypoproliferative disorder; lesion.

Homo sapiens.

WO200220032-A1.

14-MAR-2002.
10-SEP-2001; 2001WO-US028254.
08-SEP-2000; 2000US-0231212P.
(UYVO) UNIV JOHNS HOPKINS.
(UYSF-) UNIV SOUTH FLORIDA.
Yu H, Pardoll D, Jove R, Dalton W;
WPI; 2002-362218/39.
Modulating angiogenesis and an immune response in an individual, for treating a hypoxic or ischemic condition, comprises administering a compound that modulates the activity of a signal transducer and activator of transcription 3.

Disclosure; Page 83-85; 94pp; English.

The invention relates to a method of modulating angiogenesis and immune response. Method involves administering to an individual a compound that modulate the activity of signal transducer and activator of transcription 3 (Stat3). Modulating angiogenesis is useful for treating or preventing hypoxic or ischemic condition or disorder which is the result of stroke, ischaemia, coronary atherosclerosis, myocardial infarction, inflammation, tissue ischaemia in the lower extremities, infarction, trauma, vascular occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock, chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia, epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest, nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy with neovascularisation. Suppressing an immune response is useful for ameliorating a symptom of an autoimmune disease such as systemic lupus erythematosus, multiple sclerosis, insulin dependent diabetes mellitus, Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis, mixed connective tissue disease, primary biliary cirrhosis, pernicious anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo, gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis, idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and dense deposit disease. The method is useful in preventing or treating specific proliferative and oncogenic disease which includes sarcomas and carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia, fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency, hypoproliferative disorders, physical trauma, lesions and wounds. The method is also used in gene therapy. The present sequence is human protein related to angiogenesis regulation

Sequence 769 AA;

Query Match 99.2%; Score 1377; DB 5; Length 769;
Best Local Similarity 99.3%; Pred. No. 1e-114; Indels 0; Gaps 0;
Matches 269; Conservative 1; Mismatches 1;

1 RCLMEESRLQTAATAAQQGQANPTAAVTEKQOMLEOHLQVRRKRVQDLEQKKVVE 60
107 RCLMEESRLQTAATAAQQGQANPTAAVTEKQOMLEOHLQVRRKRVQDLEQKKVVE 166
QY 61 NIQDDPFRNYKTKSQGMODINGNNOSTYRQKQOLEBMTALDQMRSSYSELAGLIS 120
DB 167 NIQDDPFRNYKTKSQGMODINGNNOSTYRQKQOLEBMTALDQMRSSYSELAGLIS 226
QY 121 AMEYVQKTLTDEBLADWKRPRPACIGGPPNCLDRLEWITSLAESQQTQOQIKLEE 180
DB 227 AMEYVQKTLTDEBLADWKRPRPACIGGPPNCLDRLEWITSLAESQQTQOQIKLEE 286
QY 181 LQOKSYSGDPIVQHRPMLERIVELFRNLKMSAFVEROPCPMPHPRPRLVYKTVGQPT 240
DB 287 LQOKSYSGDPIVQHRPMLERIVELFRNLKMSAFVEROPCPMPHPRPRLVYKTVGQPT 346
QY 241 TKVRLVFPPELNYOLKIKVCIDKDSGVAA 271

Db 347 TKVRLVKKPEELNYOLKIKVCIDKDSGVAA 377

RESULT 8

AA82995 standard; protein; 770 AA.

AA82995;

25-MAR-1996 (first entry)

Mouse liver acute phase response factor.

Mouse; acute phase response factor; transcription factor; interleukin-6; signal transduction; liver; antibody; antisense; ribozyme; antiinflammatory; antitumor; hypotensive; therapy.

Mus musculus.

EP676469-A2.

11-OCT-1995.

23-MAR-1995; 95EP-00104670.

04-APR-1994; 94UP-00065825.

(KISH/) KISHIMOTO T.

Akira S, Kishimoto T;

WPI; 1995-346089/45.

N-PSDB; AAT05619.

New acute phase response factor - for developing inhibitory agents for treating diseases induced by cytokine(s) such as IL-6, e.g. inflammatory diseases.

Claim 10; Page 20-22; 31pp; English.

The sequence represents a mouse acute phase response factor (APRF), a transcription factor related to signal transduction of interleukin-6 (IL-6). The protein is encoded by a cDNA, isolated from a mouse liver cDNA library using a polymerase chain reaction product (amplified using primers derived from an IL-6-treated mouse liver peptide) as a probe. APRF-inhibitors, e.g. antibodies, antisense oligonucleotides or ribozymes, may be used to treat diseases induced by IL-6, e.g. inflammatory disease, leukemia, cancer, osteoclastia, pulmonary hypertension, etc

Sequence 770 AA;

Query Match 99.2%; Score 1377; DB 2; Length 770; Best Local Similarity 99.3%; Pred. No. 1e-114; 1; Indels 0; Gaps 0; Matches 269; Conservative 1; Mismatches 1;

1 RCIMESRLIQTATATAAAGGQANHPAAVTEKQMLEOHLOVRRKRVODLEQKKVVE 60
107 RCIMESRLIQTATATAAAGGQANHPAAVTEKQMLEOHLOVRRKRVODLEQKKVVE 166
61 NIODDDFNNKTLKSGDMODLNGNNSVTRKQMLEQMLTALDQRRSIVSELGLLS 120
167 NIODDDFNNKTLKSGDMODLNGNNSVTRKQMLEQMLTALDQRRSIVSELGLLS 226
121 AMEYVQKTLDEBELADMKRPEIACIGPPNICLDRLENNITSLAESQLOTRQIKLEE 180
227 AMEYVQKTLDEBELADMKRPEIACIGPPNICLDRLENNITSLAESQLOTRQIKLEE 286
181 LQOKSVKGDPIVOHRPMLERIVELFRNLKSAFVEROPCMRPHDRDLVKTGVQFT 240
287 LQOKSVKGDPIVOHRPMLERIVELFRNLKSAFVEROPCMRPHDRDLVKTGVQFT 346
241 TKVRLVKKPEELNYOLKIKVCIDKDSGVAA 271

Db 347 TKVRLVKKPEELNYOLKIKVCIDKDSGVAA 377

RESULT 9

AA03768 standard; protein; 770 AA.

AA03768;

11-JUN-1999 (first entry)

Human STAT3 allelic variant.

Signal Transducer and Activator of Transcription 3 (STAT3) allele; IL-6; intracellular transcription factor; interleukin-6; medicament; variant; pharmaceutical; autoimmune disease; inflammatory; human.

Homo sapiens.

EP905234-A2.

31-MAR-1999.

18-FEB-1998; 98EP-00102774.

16-SEP-1997; 97EP-00116061.

(ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.

Serlupi-Crescenzi O, Della Pietra L;

WPI; 1999-192664/17.

N-PSDB; AAX29281.

New human Signal Transducer and Activator of Transcription 3 (STAT3) allelic variant useful for treatment of autoimmune and inflammatory disease.

Claim 2; Page 9-13; 32pp; English.

The present sequence represents a predominant allelic variant of human Signal Transducer and Activator of Transcription 3 (STAT3) protein, an intracellular transcription factor which mediates IL-6 signals. The encoding sequence differs from the original published human STAT3 gene sequence (Akira et al., 1994). Host cells comprising a recombinant STAT3 DNA molecule can be used for the recombinant expression of the variant. STAT3 protein is useful as a medicament or pharmaceutical composition for treatment of autoimmune or inflammatory diseases

Sequence 770 AA;

Query Match 99.2%; Score 1377; DB 2; Length 770; Best Local Similarity 99.3%; Pred. No. 1e-114; 1; Indels 0; Gaps 0; Matches 269; Conservative 1; Mismatches 1;

1 RCIMESRLIQTATATAAAGGQANHPAAVTEKQMLEOHLOVRRKRVODLEQKKVVE 60
107 RCIMESRLIQTATATAAAGGQANHPAAVTEKQMLEOHLOVRRKRVODLEQKKVVE 166
61 NIODDDFNNKTLKSGDMODLNGNNSVTRKQMLEQMLTALDQRRSIVSELGLLS 120
167 NIODDDFNNKTLKSGDMODLNGNNSVTRKQMLEQMLTALDQRRSIVSELGLLS 226
121 AMEYVQKTLDEBELADMKRPEIACIGPPNICLDRLENNITSLAESQLOTRQIKLEE 180
227 AMEYVQKTLDEBELADMKRPEIACIGPPNICLDRLENNITSLAESQLOTRQIKLEE 286
181 LQOKSVKGDPIVOHRPMLERIVELFRNLKSAFVEROPCMRPHDRDLVKTGVQFT 240
287 LQOKSVKGDPIVOHRPMLERIVELFRNLKSAFVEROPCMRPHDRDLVKTGVQFT 346
241 TKVRLVKKPEELNYOLKIKVCIDKDSGVAA 271

Db 347 TKVRLVKEPPELNYQIKVICIDKSGDVAA 377

```

|||||
RESULT 10
AAB12377
ID AAB12377 standard; peptide; 770 AA.
XX
AC AAB12377;
XX
DT 08-NOV-2000 (first entry)
XX
DE N-terminal domain of murine STAT-3 protein.
XX
KW STAT; signal transducer and activator of transcription; crystal;
XX drug design; murine.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Region 4..9
FT /label= Alpha helix 1
FT Region 12..21
FT /label= Alpha helix 2
FT Region 19..21
FT /label= 3 (10) helix of alpha helix 2
FT Region 28..33
FT /label= Alpha helix 3
FT Region 35..40
FT /label= Alpha helix 4
FT Region 43..47
FT /label= Alpha helix 5
FT Region 50..73
FT /label= Alpha helix 6
FT Region 77..96
FT /label= Alpha helix 7
FT Region 99..119
FT /label= Alpha helix 8
XX
PN US6087478-A.
XX
PD 11-JUL-2000.
XX
PF 23-JAN-1998; 98US-00012710.
XX
PR 23-JAN-1998; 98US-00012710.
XX
PA (UNIQ) UNIV ROCKEFELLER.
XX
PI Vinkemeier U, Moarefi I, Darnell JE, Kuriyan J;
XX
DR WPI; 2000-505108/45.
XX
PT New crystals of an N-terminal fragment of a signal transducer and
PT activator of transcription that effectively diffracts x-rays, useful for
PT drug screening and development.
XX
PS Disclosure; Fig 1; 42pp; English.
XX
CC The present invention relates to a crystal of an N-terminal fragment of a
CC signal transducer and activator of transcription (STAT) protein. The
CC crystal effectively diffracts x-rays, allowing the determination of the
CC atomic coordinates of the N-terminal domain to a resolution of greater
CC than 5.0 Angstroms. The present sequence is the N-terminal domain of the
CC murine STAT 3 protein. The N-terminal domain enables STAT dimers to
CC interact and bind DNA cooperatively, a mechanism important for gene
CC activation. The crystals are useful in drug screening and development by
CC selecting a potential drug by performing rational drug design with the 3-
CC dimensional structure determined for the crystal
XX
SQ Sequence 770 AA;

```

Query Match 99.2%; Score 1377; DB 3; Length 770;

Best Local Similarity 99.3%; Pred. No. 1e-114; Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 RCLMEESRLDTAATPAAGGQGANHPATAVTEKQOMLEOHADYRKRVQDEQKMYVE 60
Db 107 RCLMEESRLDTAATPAAGGQGANHPATAVTEKQOMLEOHADYRKRVQDEQKMYVE 166
QY 61 NLQDDPFDNFKTKLSQGMQDNLGNNQSVTRQKMOLEQMLTALDQMRSSIVSELAGILS 120
Db 167 NLQDDPFDNFKTKLSQGMQDNLGNNQSVTRQKMOLEQMLTALDQMRSSIVSELAGILS 226
QY 121 AMEIVQKTLTDEELADMKRRPEIACIGPPNICDLRLNWTISLASQLOTRQOIKLEE 180
Db 227 AMEIVQKTLTDEELADMKRRPEIACIGPPNICDLRLNWTISLASQLOTRQOIKLEE 286
QY 181 LQOKVSYGDDPIVQHRPMLERIVLEFRNIMKSAFVVERQPCMPHPPRPLVIKTVGQPT 240
Db 287 LQOKVSYGDDPIVQHRPMLERIVLEFRNIMKSAFVVERQPCMPHPPRPLVIKTVGQPT 346
QY 241 TKVRLVKEPPELNYQIKVICIDKSGDVAA 271
Db 347 TKVRLVKEPPELNYQIKVICIDKSGDVAA 377

```

```

RESULT 11
AAB14652
ID AAB14652 standard; protein; 770 AA.
XX
AC AAB14652;
XX
DT 16-JUL-2002 (first entry)
XX
DE Murine STAT3 protein.
XX
KW Signal transducer and activator of transcription; STAT3;
KW drug development; drug discovery; crystal; inflammation; allergy; asthma;
KW leukæmia; anaemia; neutropænia; thrombocytopenia; cancer; obesity;
KW viral disease; growth retardation; murine.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Domain 1..130
FT /note= "Conserved N-terminal domain of the STAT family"
XX
PN US6312887-B1.
XX
PD 06-NOV-2001.
XX
PF 24-APR-2000; 2000US-00556273.
XX
PR 23-JAN-1998; 98US-00012710.
XX
PA (UNIQ) UNIV ROCKEFELLER.
XX
PI Vinkemeier U, Moarefi I, Darnell JE, Kuriyan J;
XX
DR WPI; 2002-033337/04.
XX
PT Identifying compounds that bind to signal transducer and activator of
PT transcription proteins, useful for the production of new drugs.
XX
PS Example; Col 47-50; 44pp; English.
XX
CC The invention relates to methods for detecting compounds that bind to
CC signal transducer and activator of transcription (STAT) proteins for the
CC discovery and development of new drug compounds based on the structural
CC properties of the protein crystal. The methods include: identifying a
CC compound that binds to the N-terminal domain of a STAT protein,
CC identifying a compound that enhances or diminishes the binding of the
CC dimeric STAT proteins to each other and/or their nucleic acid binding
CC site; or identifying a compound that enhances or diminishes the ability
CC of STAT protein dimers to induce the expression of a gene operably under

```

CC the control of a promoter containing at least two adjacent weak binding
 CC sites for STAT protein dimers. The methods are used for identifying new
 CC drugs. An antagonist of STAT N-terminal dimeric interactions that
 CC inhibits the binding of the STAT dimers to adjacent weak binding sites on
 CC a promoter of a gene, could be useful as drugs in the treatment of
 CC diseases, e.g. inflammation, allergy, asthma and leukaemia. On the other
 CC hand, an agonist of N-terminal dimeric interactions between STAT dimers,
 CC can be used as drugs in the treatment of diseases e.g. anaemia,
 CC neutropenia, thrombocytopaenia, cancer, obesity, viral diseases and
 CC growth retardation. The present sequence is murine STAT3 protein
 XX

Sequence 770 AA;

Query Match 99.2%; Score 1377; DB 5; Length 770;
 Best Local Similarity 99.3%; Pred. No. 1e-114;
 Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RCLMESRLIQATATAAGGAGNHPAAVTEKQMLEHLDVRRVODLEQKKVVE 60
 Db 107 RCLMESRLIQATATAAGGAGNHPAAVTEKQMLEHLDVRRVODLEQKKVVE 166
 QY 61 NLDDPDPFNKTKLSQGDMDLNGNNSVTRQKQOLEQMLTALDQRRSIVSELGLLS 120
 Db 167 NLDDPDPFNKTKLSQGDMDLNGNNSVTRQKQOLEQMLTALDQRRSIVSELGLLS 226
 QY 121 AMEYVQKTLTDEBLADMKRRPELACTGGPNNICLDLENNWITSLSASQLOTRQOIKLEE 180
 Db 227 AMEYVQKTLTDEBLADMKRRPELACTGGPNNICLDLENNWITSLSASQLOTRQOIKLEE 286
 QY 181 LOOKVSYKGDPIVQHRPMLERIVELFRNLMSAFVEROPCMHPMDRPLVTKTGQFT 240
 Db 287 LOOKVSYKGDPIVQHRPMLERIVELFRNLMSAFVEROPCMHPMDRPLVTKTGQFT 346
 QY 241 TKVRLLVKPEELNYQLKIKVCIDKSGDVAA 271
 Db 347 TKVRLLVKPEELNYQLKIKVCIDKSGDVAA 377

RESULT 12

ABG69497
 ID ABG69497 standard; protein; 770 AA.

XX ABG69497;

XX 21-OCT-2002 (first entry)

XX Human bait protein STAT3.

XX Human; Yeast two-hybrid assay; adipocyte; bait protein; NIDDM;

XX non-insulin diabetes mellitus; obesity; selected interacting domain; SID;

XX protein-protein interaction map; PIM; anorectic; metabolic disorder.

XX Homo sapiens.

XX MO200253726-A2.

XX 11-UTL-2002.

XX 28-DEC-2001; 2001WO-EP015423.

XX 02-JAN-2001; 2001US-0259377P.

XX (HYBR-) HYBRIGENICS.

XX (CNRS) CENT NAT RECH SCI.

XX Legrain P, Marullo S, Jockers R;

XX WPI: 2002-583612/62.

XX N-PSDB; AB851033.

XX Novel complex of protein-protein interactions in adipocyte cells for
 PT identifying compounds that modulate the protein-protein interactions and
 PT useful for treating obesity and metabolic disorders.

XX Claim 1; Page 54; 125pp; English.

PS The invention relates to a complex of protein-protein interactions
 XX (forming a protein-protein interaction map, PIM) in adipocyte cells as
 CC defined in the specification, or polynucleotides in adipocytes encoding the
 CC for the polypeptides. Also included are a recombinant cell expressing the
 CC interacting polypeptides and a method of selecting a modulating compound
 CC in adipocyte cells, by culturing a recombinant host cell on a selective
 CC medium containing a modulating compound and a reporter gene the
 CC expression of which is toxic for the recombinant host cell which is
 CC transformed with two vectors, where the first vector comprises a
 CC polynucleotide encoding a first hybrid polypeptide and DNA binding domain
 CC and the second vector comprising a polynucleotide encoding a second
 CC hybrid polypeptide and an activating domain that activates the toxic
 CC reporter gene, when the first and second hybrid polypeptides interact and
 CC selecting the modulating compound which inhibits the growth of the
 CC recombinant host cell (i.e. using the yeast two-hybrid system). The
 CC complexes are useful for identifying compounds that modulate the protein-
 CC protein interactions and useful for treating obesity and metabolic
 CC disorders e.g. non-insulin dependent diabetes mellitus, NIDDM. The
 CC compound isolated by the method is useful for treating and preventing
 CC obesity or metabolic diseases. The interactions between the proteins of
 CC the complex further define a set of selected interacting domains, SID.
 CC The present sequence represents a member of the protein complex of the
 CC invention, used as the bait protein in the yeast two-hybrid assay
 XX

Sequence 770 AA;

Query Match 99.2%; Score 1377; DB 5; Length 770;
 Best Local Similarity 99.3%; Pred. No. 1e-114;
 Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RCLMESRLIQATATAAGGAGNHPAAVTEKQMLEHLDVRRVODLEQKKVVE 60
 Db 107 RCLMESRLIQATATAAGGAGNHPAAVTEKQMLEHLDVRRVODLEQKKVVE 166
 QY 61 NLDDPDPFNKTKLSQGDMDLNGNNSVTRQKQOLEQMLTALDQRRSIVSELGLLS 120
 Db 167 NLDDPDPFNKTKLSQGDMDLNGNNSVTRQKQOLEQMLTALDQRRSIVSELGLLS 226
 QY 121 AMEYVQKTLTDEBLADMKRRPELACTGGPNNICLDLENNWITSLSASQLOTRQOIKLEE 180
 Db 227 AMEYVQKTLTDEBLADMKRRPELACTGGPNNICLDLENNWITSLSASQLOTRQOIKLEE 286
 QY 181 LOOKVSYKGDPIVQHRPMLERIVELFRNLMSAFVEROPCMHPMDRPLVTKTGQFT 240
 Db 287 LOOKVSYKGDPIVQHRPMLERIVELFRNLMSAFVEROPCMHPMDRPLVTKTGQFT 346
 QY 241 TKVRLLVKPEELNYQLKIKVCIDKSGDVAA 271
 Db 347 TKVRLLVKPEELNYQLKIKVCIDKSGDVAA 377

RESULT 13

ABU10476
 ID ABU10476 standard; protein; 770 AA.

XX ABU10476;

XX 06-AUG-2003 (first entry)

XX Mouse STAT3 protein.

XX Mouse; signal transducer and activator of transcription; drug design;

XX drug screening; STAT-STAT dimer interaction; STAT3.

XX Mus sp.

XX Key

XX Region

XX Region

Location/Qualifiers

4..9

/label= alpha_helix_1

12..21

FT	/label= alpha_helix_2
FT	/note= "Residues 19-21 form a 3 helix"
FT	28. .33
FT	/label= alpha_helix_3
FT	35. .40
FT	/label= alpha_helix_4
FT	43. .47
FT	/label= alpha_helix_5
FT	50. .73
FT	/label= alpha_helix_6
FT	/note= "Residues 57, 61, 64, 68 and 71 contribute to
FT	packing of the coiled-coil"
FT	77. .96
FT	/label= alpha_helix_7
FT	/note= "Residues 79, 83, 86, 90 and 94 contribute to
FT	packing of the coiled-coil"
FT	99. .119
FT	/label= alpha_helix_8
XX	
PN	US2003003563-A1.
PD	
XX	02-JAN-2003.
XX	
PF	19-OCT-2001; 2001US-00045792.
XX	
PR	23-JAN-1998; 98US-00012710.
PR	24-APR-2000; 2000US-00556273.
XX	
PA	(VINK/) VINKEMEIER U.
PA	(MOAR/) MOAREFI I.
PA	(DARN/) DARNELL J E.
PA	(KURI/) KURIYAN J.
PI	Vinkemeier U, Moarefi I, Darnell JE, Kuriyan J;
XX	
DR	WPI; 2003-447354/42.
PT	
PT	New crystal having an N-terminal domain of a STAT protein performing X-
PT	ray crystallographic studies, useful for screening drugs that enhance or
XX	inhibit STAT-STAT dimer interactions.
XX	
PS	Disclosure; Page 25-26; 46pp; English.
CC	
CC	The invention relates to a crystal of an N-terminal domain of signal
CC	transducer and activator of transcription (STAT) protein, where the
CC	crystal effectively diffracts X-rays for the determination of the atomic
CC	coordinates of the N-terminal domain of the STAT protein to a resolution
CC	of greater than 5.0 Angstrom. The methods and compositions are useful for
CC	the design and screening of drugs that enhance or inhibit STAT-STAT dimer
CC	interactions. The present sequence represents the amino acid sequence of
CC	mouse STAT3 protein
XX	
SQ	Sequence 770 AA;

Query Match	99.2%	Score 1377	DB 6	Length 770
Best Local Similarity	99.3%	Pred. No. 1e-114		
Matches 269	Conservative 1	Mismatches 1	Indels 0	Gaps 0
QY	1	RCIMESRLLOQTAAATAAQQGGGGAHNPPTAAVYTEKQOMLEQHLQDVAKRVODLEQKMKVVE	60	
DB	107	RCIMESRLLOQTAAATAAQQGGGGAHNPPTAAVYTEKQOMLEQHLQDVAKRVODLEQKMKVVE	166	
QY	61	NIQDDPDPHYKTKLKSQGMQDINGNNOSTRTQKMOQLEOMLTALOMRSTSYSELAGLIS	120	
DB	167	NIQDDPDPHYKTKLKSQGMQDINGNNOSTRTQKMOQLEOMLTALDDMRSTSYSELAGLIS	226	
QY	121	AMEYVQKTLTDEELADWKRREPIACIGCPNLCIDRLNEMWITSLASQQRQOQIKLEE	180	
DB	227	AMEYVQKTLTDEELADWKRROQACIGCPNLCIDRLNEMWITSLASQQRQOQIKLEE	286	
QY	181	LOOKYSYKDDPIVOHNPMLPEERIVELFRNIMSAFAVERQPCMPHDPDPVYKTVQPT	240	
DB	287	LOOKYSYKDDPIVOHNPMLPEERIVELFRNIMSAFAVERQPCMPHDPDPVYKTVQPT	346	

Cy 241 TKVRLYKFPPELANTQLKVICIDSGDVAA 271
Db 347 TKVRLYKFPPELANTQLKVICIDSGDVAA 377

RESULT 14
ADN04365
ID ADN04365 standard; protein; 770 AA.
XX AC ADN04365;
XX DT 01-JUL-2004 (first entry)
XX DE Antipsoriatic protein sequence #377.
XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.
XX OS Homo sapiens.
XX PN WO2004028479-A2.
XX PD 08-APR-2004.
XX PF 25-SEP-2003; 2003WO-USO30907.
XX PR 25-SEP-2002; 2002US-041406P.
XX PA (GETH) GENENTECH INC.
XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI,
XX PU TD;
XX WPI; 2004-305105/28.
XX DR N-PSDB; ADN04364.
PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
PS Claim 9; SEQ ID NO 759; 3069pp; English.
CC The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polypeptides of the invention.

Sequence 770 AA;

Query Match	Similarity	Score	ID	Length
Best Local	99.2%	1377	DB 8	770
Matches	269	Conservative	1	Mismatches 1, Indels 0, Gaps 0

Qy	1	RLWESRLTQTATTAAGGGAAHPHTAAVTEKQMLRQHLQDYKRKVQDLEQMKVYE	60
Db	107	RLWBSRLTQTATTAAGGGAAHPHTAAVTEKQMLEQLQDYKRKVQDLEQMKVYE	166
Qy	61	NIQDFFDNVYKTKSQGDMODLNGNNSQVTRQKMQOLEQMTALDQWRSIVSELAGLLS	120
Db	167	NIQDFFDNVYKTKSQGDMODLNGNNSQVTRQKMQOLEQMTALDQWRSIVSELAGLLS	226
Qy	121	AMEYQKTLTDBELADWRRPEIACIGSPNII CDRLFNMTTSAESQLOTRQOKKLEE	180
Db	227	AMEYQKTLTDBELADWRRPEIACIGSPNII CDRLFNMTTSAESQLOTRQOKKLEE	286
Qy	181	LOQKVSXGDDPIVQHRPMLERIVELFNNLMKSAFVERROPMPMHPDRPIVTKGVQFT	240
Db	287	LOQKVSXGDDPIVQHRPMLERIVELFNNLMKSAFVERROPMPMHPDRPIVTKGVQFT	346
Qy	241	TKVRLVAKPELNIYOLKIKVICDKDSGVAA	271
Db	347	TKVRLVAKPELNIYOLKIKVICDKDSGVAA	377

RESULT 15
AAB58442
ID AAB58442 standard; protein; 793 AA.
XX
AC AAB58442;
DT 14-MAR-2001 (first entry)
XX
DE Lung cancer associated polypeptide sequence SEQ ID 780.
XX
KW Human; lung cancer associated protein; neuroprotective; cyrostatic;
KW cardioactive; immunomodulatory; muscular active; vulnary;
KW gastrointestinal; nephrotoxic; anti-infective; gynecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease.
XX
OS Homo sapiens.
XX
EN W020005180-A2.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000MO-US005918.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI (ROSE/) ROSEN C A.
XX
PI Ruben SM;
XX
DR WPI; 2000-587514/55.
XX
DR N-PSDB; AAF18318.
XX
PT Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer.
XX
PS Claim 11; Page 1310-1313; 1425pp; English.
XX
CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cyrostatic; cardioactive;
CC immunomodulatory; muscular active general; vulnary; gastrointestinal
CC general; nephrotoxic; anti-infective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the protein
CC or polynucleotide sequences. The lung cancer associated polynucleotide
CC sequences may be used for detection of lung cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The proteins may be used to treat disorders such as
CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
CC cardiovascular, renal, and proliferative disorders. The proteins may also
CC be used in the treatment of wounds and infectious diseases.
CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are
CC used in the course of the invention for the identification and
CC characterisation of the polynucleotide and protein sequences
XX
SQ Sequence 793 AA;

Query Match 99.2%; Score 1377; DB 3; Length 793;
Best Local Similarity 99.3%; Pred. No. 1e-114;
Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RCLMESRLIQATATAAOCAGANHPAAVTEKQOMLEOHLDVRRKRVODLEOKMKVVE 60
DB 130 RCLMESRLIQATATAAOCAGANHPAAVTEKQOMLEOHLDVRRKRVODLEOKMKVVE 189
QY 61 NLQDDDFPNYKTLKSGDMQDLNGNNSVTRQKQOLEQMLTALDQMRSSIVSELGLLS 120
DB 190 NLQDDDFPNYKTLKSGDMQDLNGNNSVTRQKQOLEQMLTALDQMRSSIVSELGLLS 249

QY 121 AMEYVQKTLNDEELADWKRREPEIACIGPPNLCIDRLNMTSLAESQLOTRQIKKLE 180
DB 250 AMEYVQKTLNDEELADWKRREPEIACIGPPNLCIDRLNMTSLAESQLOTRQIKKLE 309
QY 181 LQKVSYSKDPPIVOHRPMLERIVLEFRNLKSAFVEROPCMHPDRPIVITGVQFT 240
DB 310 LQKVSYSKDPPIVOHRPMLERIVLEFRNLKSAFVEROPCMHPDRPIVITGVQFT 369
QY 241 TKVRLVKKPEELNLYOLKIKVCIDKSGDVAA 271
DB 370 TKVRLVKKPEELNLYOLKIKVCIDKSGDVAA 400

Search completed: December 2, 2004, 23:49:06
Job time : 161 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2004, 23:35:28 ; Search time 193 Seconds

(without alignments)
807.909 Million cell updates/sec

Title: US-10-090-185-9

Perfect score: 1388
Sequence: 1 RCLMEBSRLQTAATAAQQG.....LNYQIKIKVICIDKSGDVAA 271

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1377	99.2	722	AAH19168	Aah19168 mus muscu
2	1377	99.2	770	STAB3_HUMAN	P40763 homo sapien
3	1377	99.2	770	STAB3_MOUSE	P42227 mus musculu
4	1374	99.0	770	STAB3_RAT	P52631 rattus norv
5	1370	98.7	770	STAB3_BOVIN	P61635 bos taurus
6	1352	97.4	771	O6DV79	O6dv79 gallus gall
7	1304	93.9	769	O9PVX8	O9pvx8 xenopus lae
8	1286	92.7	766	O7ZKX3	O7zxx3 xenopus lae
9	1191.5	85.8	786	O6WV46	O6wv46 brachydant
10	1191.5	85.8	786	AAH68320	Aah68320 brachydant
11	1191.5	85.8	786	O93599	O93599 brachydant
12	1191.5	85.8	806	O6GUE7	O6gue7 oryzias lat
13	1182.5	85.2	765	O6GUE7	O6gue7 oryzias lat
14	1182.5	85.2	785	O90Y16	O90y16 tetrarodon f
15	1165.5	84.0	764	O13133	O13133 oncorhynch
16	1147.5	82.7	767	O9N145	O9n145 macaca mula
17	754	54.3	163	O8GNO	O8gno xenopus lae
18	654	47.8	751	O6P6Q7	O6p6q7 rattus norv
19	654	47.1	712	AAH62079	Aah62079 rattus no
20	654	47.1	712	O9QXK0	O9qxx0 rattus norv
21	654	47.1	749	O9QXK0	O9qxx0 rattus norv
22	652	47.0	749	O9QXK4	O9qxx4 mus musculu
23	651	46.9	712	O8C3V4	O8c3v4 mus musculu
24	651	46.9	749	O9D323	O9d323 mus musculu
25	651	46.9	755	O8C8M3	O8c8m3 mus musculu
26	651	46.9	712	AAH35905	Aah35905 homo sapi
27	647	46.6	750	STAB1_HUMAN	P42224 homo sapien
28	647	46.6	757	O764M5	O764m5 sus scrofa
29	646	46.5	757	BAD06318	Bad06318 sus scrofa
30	646	46.5	749	STAB1_MOUSE	P42225 mus musculu
31	619	44.6	749	STAB1_MOUSE	P42225 mus musculu

32	606	43.7	754	2	O13131	O13131 oncorhynch
33	602	43.4	718	2	O801Y2	O801y2 carassius a
34	596	42.9	758	2	O90Y17	O90y17 tetrarodon f
35	595.5	42.9	528	1	O8F08	O8f08 brachydant
36	595.5	42.9	748	1	STAB3_HUMAN	STAB3 homo sapien
37	587.5	42.3	657	2	O8AW24	O8aw24 brachydant
38	585	42.1	754	2	O13132	O13132 oncorhynch
39	580.5	41.8	749	2	O93598	O93598 brachydant
40	580.5	41.8	749	2	O6P943	O6p943 brachydant
41	580.5	41.8	749	2	AAH60933	Aah60933 brachydant
42	579	41.7	749	1	STAB3_MOUSE	STAB3 mus musculu
43	578.5	41.7	651	2	O7Z253	O7z253 brachydant
44	578.5	41.7	667	2	O8AW20	O8aw20 brachydant
45	578	41.6	652	2	O7Z277	O7z277 brachydant

ALIGNMENTS

RESULT 1										
AAH19168	PRELIMINARY;	PRT;	722 AA.							
AC	AAH19168;									
DT	02-MAR-2004 (TREMBLrel. 27, Created)									
DT	02-MAR-2004 (TREMBLrel. 27, Last sequence update)									
DT	02-MAR-2004 (TREMBLrel. 27, Last annotation update)									
DE	Signal transducer and activator of transcription 3.									
OS	Mus musculus (Mouse).									
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.									
OX	NCBI_TaxID=10090;									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=FVB/N; TISSUE=Salivary gland;									
RX	MEDLINE=22386257; PubMed=12477932;									
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,									
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,									
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,									
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,									
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,									
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,									
RA	Brownstein M.J., Uedin T.B., Toshitoki S., Carrinci P., Prange C.J.,									
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,									
RA	Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,									
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halys S.W.,									
RA	Villalona D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,									
RA	Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,									
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,									
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,									
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,									
RA	Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,									
RA	Jones S.J., Marra M.A.;									
RT	"Generation and initial analysis of more than 15,000 full-length human									
RL	and mouse cDNA sequences."									
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).									
RN	[2]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=FVB/N; TISSUE=Salivary gland;									
RC	Strausberg R.;									
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBSJ databases.									
DR	EMBL; BC019166; AAH19168.1; -									
SO	SEQUENCE 722 AA; 83126 MW; 09226A697966D947 CRC64;									
QY	Query Match									
	Best Local Similarity	99.2%;	Score 1377;	DB 2;	Length 722;					
	Matches	269;	Conservative	1;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	RCLMEBSRLQTAATAAQQG	30ANHPAAVTEKQMLEBHLIDVRRVODILEQKRYVE	60						
DB	107	RCLMEBSRLQTAATAAQQG	30ANHPAAVTEKQMLEBHLIDVRRVODILEQKRYVE	166						
QY	61	NLQDDPDPNYTKISQGMOLNNGNSVTRQXMOQLPOMLTALQPMRSTYSELAGILS	120							

```

Db      167 NIODDFENYKTKISQGDMDLNGNNOSVTRQMOLEQMLTALDQMRSSIVSEIAGLIS 226
Qy      121 AMEYVQKTLDEBLADMKRPEREACIGSPENICLDLEMMITSLAESQLOTRQOIKLEE 180
Db      227 AMEYVQKTLDEBLADMKRROQIACIGSPENICLDLEMMITSLAESQLOTRQOIKLEE 286
Qy      181 LOOKKYSKGDPIYQHRPMEERIVELPRNIMKSAFVEROPCPMPHDPRLVITGVQFT 240
Db      287 LOOKKYSKGDPIYQHRPMEERIVELPRNIMKSAFVEROPCPMPHDPRLVITGVQFT 346
Qy      241 TKVRLVKEPELNYOLKIKVCIKDKSGDVAA 271
Db      347 TKVRLVKEPELNYOLKIKVCIKDKSGDVAA 377

RESULT 2
STA3 HUMAN STANDARD; PRT; 770 AA.
ID STA3 HUMAN
AC P40763; O14916; Q9BWS4;
DT 01-FEB-1995 (Rel. 31, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Signal transducer and activator of transcription 3 (Acute-phase response factor).
GN Name=STAT3; Synonyms=APRF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Placenta;
RX MEDLINE=94208062; PubMed=7512451;
RA Akita S., Nishio Y., Inoue M., Wang X.-J., Wei S., Matsusaka T., Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related transcription factor involved in the gp130-mediated signaling pathway.";
RL Cell 77:63-71(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=9829260; PubMed=9630560;
RA Della Pietra L., Bressan A., Pezzotti A., Serlupi-Crescenzi O.;
RT "Highly conserved amino-acid sequence between murine STAT3 and a revised human STAT3 sequence.";
RL Gene 213:119-124(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM DEL-701), AND VARIANT ILE-143.
RA Rieder M.J., Daniels R.L., da Ponte S.H., Hastings N.C., Ahearn M.O., Rajkumar N., Yi Q., Nickerson D.A.;
RT "SeacLESNP: NHBI HU6682 program for genomic applications, UW-FHCR, Seattle, WA (URL: http://pga.gs.washington.edu).";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND DEL-701).
RC TISSUE=Kidney and Pancreas;
RX MEDLINE=22388257; PubMed=1247932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G., Schler G.D., Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loughlano N.A., Peters G.J., Abramson R.D., Mullighy S.J., Bosk S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W., Villalón D.K., Muzny D.M., Sodergren E.U., Lu X., Gibbs R.A., Rahy U., Helton E., Kettman W., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,

```

```

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SEQUENCE OF 564-704 FROM N.A.
RC TISSUE=Liver;
RA Della Pietra L., Bressan A., Pezzotti A.R., Serlupi-Crescenzi O.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP PHOSPHORYLATION ON SERINE.
RX MEDLINE=95215843; PubMed=7010321;
RA Zhang X., Blenis J., Li H.-C., Schindler C., Chen-Kiang S.;
RT "Requirement of serine phosphorylation for formation of STAT-promoter complexes.";
RL Science 267:1990-1994(1995).
CC -1- FUNCTION: Transcription factor that binds to the interleukin-6 (IL-6)-responsive elements identified in the promoters of various acute-phase protein genes.
CC -1- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC -1- SUBUNIT: Forms a homodimer or a heterodimer with a related family member (at least STAT1).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus in response to phosphorylation.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P40763-1; Sequence=Displayed;
CC Name=Del-701;
CC IsoId=P40763-2; Sequence=VSP_010474;
CC -1- TISSUE SPECIFICITY: Heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas.
CC -1- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF, LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation is important for the formation of stable DNA-binding STAT3 homodimers and maximal transcriptional activity.
CC -1- SIMILARITY: Belongs to the transcription factor STAT family.
CC -1- SIMILARITY: Contains 1 SH2 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L29277; AA58374.1; -
DR EMBL; AJ012463; CA010032.1; -
DR EMBL; AY572796; AA566986.1; -
DR EMBL; BC000627; AAH00627.1; -
DR EMBL; BC014482; AAH14482.1; -
DR EMBL; AF029311; AAB84254.1; -
DR PIR; A54444; A54444.
DR HSSP; P42227; 1BG1.
DR TRANSFAC; T01493; -
DR Genew; HGNC:11364; STAT3.
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0005062; F:hematopoietin/interferon-class (D200-domain); TAS.
DR GO; GO:0003700; F:transcription factor activity; TAS.
DR GO; GO:0006928; P:cell motility; TAS.
DR GO; GO:0007259; P:JAK-STAT cascade; TAS.
DR GO; GO:0000122; P:negative regulation of transcription from P.; TAS.
DR GO; GO:0007399; P:neurogenesis; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR008367; P53_like_DNA_Bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_Dbd; 1.

```

PFam: PRO2865; STAT_int; 1.
 DR PROSITE; PS50001; SH2; 1.
 KW Activator; Alternative splicing; DNA-binding; Nuclear protein;
 KM Phosphorylation; Polymorphism; SH2 domain; Transcription regulation.
 FT DOMAIN 580 670 SH2.
 FT MOD_RES 705 705 Phosphotyrosine (by JAK) (By similarity).
 FT MOD_RES 727 727 Phosphoserine (by similarity).
 FT VARSPLOC 701 701 Missing (in isoform Del-701).
 FT VARIANT 32 32 Q -> K (in dbSNP:1803125).
 FT VARIANT 143 143 M -> I.
 FT VARIANT 143 143 /FTID=VAR_018683.
 FT CONFLICT 288 288 P -> H (in Ref. 1).
 FT CONFLICT 460 460 Q -> S (in Ref. 1).
 FT CONFLICT 548 548 K -> N (in Ref. 1).
 FT CONFLICT 561 561 F -> Y (in Ref. 1).
 FT CONFLICT 667 667 V -> L (in Ref. 1).
 FT CONFLICT 730 730 T -> A (in Ref. 1).
 SQ SEQUENCE 770 AA; 88067 MW; 6C00632211C8012D CRC64;

Query Match 99.2%; Score 1377; DB 1; Length 770;
 Best Local Similarity 99.3%; Pred. No. 2e-77;
 Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RCLWESRLQTAATAAQQGGGQANPTAAVTEKQOMLEQHLQDVKKVQDLEQKRYVE 60
 DB 107 RCLWESRLQTAATAAQQGGGQANPTAAVTEKQOMLEQHLQDVKKVQDLEQKRYVE 166
 QY 61 NLQDQFDNNYKTKSGQGMQDNLGNNSVTRQKMOQLQMLTALQMRKSYSELAGLS 120
 DB 167 NLQDQFDNNYKTKSGQGMQDNLGNNSVTRQKMOQLQMLTALQMRKSYSELAGLS 226
 QY 121 AMEYVQKLTDBELADMKRRPEIACIGPPNICDRLNMTSLAESQLQTRQKIKLEB 180
 DB 227 AMEYVQKLTDBELADMKRRQIACIGPPNICDRLNMTSLAESQLQTRQKIKLEB 286
 QY 181 LQQRVYKGDPIVQHPMLKEEIVLFLNLMKSAFVERQCPMPHDPRLVIKTVQFT 240
 DB 287 LQQRVYKGDPIVQHPMLKEEIVLFLNLMKSAFVERQCPMPHDPRLVIKTVQFT 346
 QY 241 TKVRLVKEPELNYQTKIKVCIDKSGDVAA 271
 DB 347 TKVRLVKEPELNYQTKIKVCIDKSGDVAA 377

RESULT 3

ID STAS_MOUSE STANDARD; PRT; 770 AA.
 AC P42227;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Signal transducer and activator of transcription 3 (Acute-phase response factor).
 GN Name=Stat3; Synonyms=Apf1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM STAT3A), AND SEQUENCE OF 154-158; 181-185
 RP AND 632-640.
 RC STRAIN=BALB/c; TISSUE=Liver;
 RX MEDLINE=94208062; PubMed=7512451;
 RA Akira S., Niepio Y., Inoue M., Wang X.J., Shi W., Matsusaka T.,
 RA Yoshida K., Sudo T., Naruto M., Kishimoto T.,
 RT "Molecular cloning of APRF, a novel IFN-gamma-stimulated gene factor 3 p91-
 RT related transcription factor involved in the gp130-mediated signaling
 RT pathway.";
 RL Cell 77:63-71(1994).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM STAT3A).

RC TISSUE=Thymus;
 RA MEDLINE=94188718; PubMed=8140422;
 RX Zhong Z., Wen Z., Darnell J.E. Jr.;
 RT "Stat3: a STAT family member activated by tyrosine phosphorylation in
 RT response to epidermal growth factor and interleukin-6.";
 RL Science 264:95-98(1994).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM STAT3A).
 RC TISSUE=Brain;
 RX MEDLINE=95014185; PubMed=7523373;
 RA Raz R., Durbin J.E., Levy D.E.;
 RT "Acute phase response factor and additional members of the interferon-
 RT cytokines, interferons, and growth factors.";
 RL J. Biol. Chem. 269:24391-24395(1994).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM STAT3B).
 RC STRAIN=BALB/c; and C57BL/6; TISSUE=Liver;
 RX MEDLINE=96016116; PubMed=7568080;
 RA Schaefer T.S., Sanders L.K., Nathans D.;
 RT "Cooperative transcriptional activity of Jun and Stat3 beta, a short
 RT form of Stat3.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:9097-9101(1995).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM STAT3A).
 RC STRAIN=129/SvJ;
 RX PubMed=11161808; DOI=10.1006/geno.2000.6433;
 RA Miyoshi K., Cui Y., Riedlinger G., Robinson P., Lehoczy J., Zon L.,
 RT Oka T., Dewar K., Hennighausen L.;
 RT "Structure of the mouse Stat 3/5 locus: evolution from Drosophila to
 RT zebrafish to mouse.";
 RL Genomics 71:150-155(2001).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM STAT3A).
 RC STRAIN=C57BL/6J, and NOD/LtJ;
 RA Davoodi-Semlomi A., She J.-X.;
 RT "A mutant Stat3 with weaker DNA binding defines a key defective
 RT pathway in non-obese diabetic (NOD) mice.";
 RL Submitted (MAY-2003) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORM STAT3A).
 RC STRAIN=FVB/N; TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Sherman C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usetin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullah S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywicki M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8]
 RP PHOSPHORYLATION SITE SER-727, AND MUTAGENESIS.
 RX MEDLINE=95354205; PubMed=7543024;
 RA Wen Z., Zhong Z., Darnell J.E. Jr.;
 RT "Maximal activation of transcription by Stat1 and Stat3 requires both
 RT tyrosine and serine phosphorylation.";
 RL Cell 82:241-250(1995).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 136-716.
 RX MEDLINE=98334373; PubMed=9671298;

RA Becker S., Groner B., Mueller C.W.;
 RT "Three-dimensional structure of the Stat3beta homodimer bound to
 RT DNA.";
 RL Nature 394:145-151 (1998).
 CC -I- FUNCTION: Transcription factor that binds to the interleukin-6
 CC (IL-6)-responsive elements identified in the promoters of various
 CC acute-phase protein genes. STAT3B interacts with the N-terminal
 CC part of JUN to activate such promoters in a cooperative way.
 CC -I- PATHWAY: Involved in the gp130-mediated signaling pathway.
 CC -I- SUBUNIT: Forms a homodimer or a heterodimer with a related family
 CC member (at least STAT1) (By similarity).
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
 CC in response to phosphorylation.
 CC -I- ALTERNATIVE PRODUCTS:
 CC Name=Stat3A;
 CC IsoId=P42227-1; Sequence=Displayed;
 CC Name=Stat3B;
 CC IsoId=P42227-2; Sequence=VSP_006287;
 CC Name=Del-701;
 CC IsoId=P42227-3; Sequence=VSP_010475;
 CC -I- TISSUE SPECIFICITY: STAT3A is seen in the liver, spleen, and
 CC kidney. STAT3B is also detected in the liver, although in a much
 CC less abundant manner.
 CC -I- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
 CC LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
 CC is important for the formation of stable DNA-binding STAT3
 CC homodimers and maximal transcriptional activity (By similarity).
 CC -I- SIMILARITY: Belongs to the transcription factor STAT family.
 CC -I- SIMILARITY: Contains 1 SH2 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, L29278; AAA7254.1; -;
 DR EMBL, U06922; AAA19452.1; -;
 DR EMBL, U08378; AAA56668.1; -;
 DR EMBL, U30709; AAC52612.1; -;
 DR EMBL, AF246978; AAL59017.1; -;
 DR EMBL, AY299489; AAQ75418.1; -;
 DR EMBL, AY299490; AAQ75419.1; -;
 DR EMBL, BC003806; AAH03806.1; -;
 DR PIR, I49508; I49508.
 DR PDB, 1BG1; X-ray; A=1-722.
 DR TRANSFAC; T01574; -;
 DR MGD; MGI:103038; Stat3.
 DR GO; GO:0005737; Cytoplasm; IDA.
 DR GO; GO:0005634; Cnucleus; IDA.
 DR GO; GO:0005886; Cplasma membrane; IDA.
 DR GO; GO:0003677; F-DNA binding; IDA.
 DR GO; GO:0005515; F-protein binding; IPI.
 DR GO; GO:0015653; F-transcriptional activator activity; IDA.
 DR GO; GO:0007259; P-JAK-STAT cascade; IDA.
 DR GO; GO:0006357; P-regulation of transcription from Pol II pro. .; IDA.
 DR InterPro; IPR008967; P53_like_DNA_bnd.
 DR InterPro; IPR000980; SH2.
 DR InterPro; IPR001217; STAT.
 DR Pfam; PF00017; SH2; 1.
 DR Pfam; PF01017; STAT_alpha; 1.
 DR Pfam; PF02864; STAT_bind; 1.
 DR Pfam; PF02865; STAT_int; 1.
 DR PROSITE; PSS0001; SH2; 1.
 DR 3D-structure; Activator; Acute phase; Alternative splicing;
 DR Direct protein sequencing; DNA-binding; Nuclear protein;
 DR Phosphorylation; SH2 domain; Transcription regulation.
 DR DOMAIN 580 670
 DR MOD_RES 705 705 Phosphotyrosine (by JAK) (By similarity).
 DR MOD_RES 727 727 Phosphoserine.

FT	VARSPLIC	716	770	TTCSNTIDLPMSPRTLSIMOFGNNGEAPSGQFESLT FMDLTSBECATSPM -> FIDAVMK (in isoform Stat3B). /FTid=VSP_006287. Missing (in isoform Del-701). S->A: Decreased transcriptional activation. B -> K (in Ref. 2). S -> T (in Ref. 2 and 4). M -> I (in Ref. 1).
FT	VARSPLIC	701	701	
FT	MUTAGEN	727	727	
FT	CONFLICT	16	16	
FT	CONFLICT	25	25	
FT	CONFLICT	394	394	
FT	HELIX	139	180	
FT	TURN	181	182	
FT	TURN	197	198	
FT	HELIX	199	237	
FT	TURN	238	238	
FT	HELIX	239	251	
FT	TURN	252	253	
FT	HELIX	261	280	
FT	TURN	294	295	
FT	TURN	297	301	
FT	HELIX	302	320	
FT	STRAND	321	328	
FT	TURN	330	331	
FT	TURN	332	334	
FT	TURN	336	337	
FT	STRAND	341	340	
FT	TURN	341	342	
FT	STRAND	345	351	

Query Match 99.2%; Score 1377; DB 1; Length 770;
 Best Local Similarity 99.3%; Pred. No. 2e-77; 1; Indels 0; Gaps 0;
 Matches 269; Conservative 1; Mismatches 1;
 Db 1 RCLMESRLQTAATPAAGGQANPTAAVTEKQMLEQHIDPVKRYQDLEQKKAYE 60
 107 RCLMESRLQTAATPAAGGQANPTAAVTEKQMLEQHIDPVKRYQDLEQKKAYE 166
 QY 61 NIQDDPFENYKTLKSGDMQDNGNNSVTRQKQOLEQMTALDQMRISVSELAGLLS 120
 Db 167 NIQDDPFENYKTLKSGDMQDNGNNSVTRQKQOLEQMTALDQMRISVSELAGLLS 226
 QY 121 AMEVQKLTDEELADMKRRPFIACIGPPNLCIRLEWITSLSAQOTROQIKKEE 180
 Db 227 AMEVQKLTDEELADMKRRQIACIGPPNLCIRLEWITSLSAQOTROQIKKEE 286
 QY 181 LQOKSYKQDPVQHRPMLERIVELFRNLMSAFVERQPCPMHPDRLVIKTVGQPT 240
 Db 287 LQOKSYKQDPVQHRPMLERIVELFRNLMSAFVERQPCPMHPDRLVIKTVGQPT 346
 QY 241 TVRRLIVKPELNYQIKIVCIDKSGDVAA 271
 Db 347 TVRRLIVKPELNYQIKIVCIDKSGDVAA 377
 RESULT 4
 ID STAT3 RAT STANDARD; PRT; 770 AA.
 AC P52631;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Signal transducer and activator of transcription 3.
 GN Name=Stat3;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OK NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=96102059; PubMed=8530402;
 RA Ripberger J.A., Filtz S., Richter K., Hooke G.M., Lottepeich F.,

RA Fey G.H.;
 RT "Transcription factors Stat3 and Stat5b are present in rat liver
 RT nuclei late in an acute phase response and bind interleukin-6 response
 RT elements.";
 RL J. Biol. Chem. 270:29998-30006(1995).
 CC
 CC -1- FUNCTION: Transcription factor that binds to the interleukin-6
 CC (IL-6)-responsive elements identified in the promoters of various
 CC acute-phase protein genes.
 CC -1- PATHWAY: Involved in the gp130-mediated signaling pathway.
 CC -1- SUBUNIT: Forms a homodimer or a heterodimer with a related family
 CC member (at least STAT1) (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
 CC in response to phosphorylation (By similarity).
 CC -1- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
 CC ILF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
 CC is important for the formation of stable DNA-binding STAT3
 CC homodimers and maximal transcriptional activity (By similarity).
 CC -1- SIMILARITY: Belongs to the transcriptional factor STAT family.
 CC -1- SIMILARITY: Contains 1 SH2 domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X91810; CAA62920.1; -.
 DR HSSP: P42227; 18G1.
 DR RGD: 3772; Stat3.
 DR InterPro: IPR008967; P53_like_DNA_bnd.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF01017; STAT_alpha; 1.
 DR Pfam: PF02864; STAT_bind; 1.
 DR Pfam: PF02865; STAT_int; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS00001; SH2; 1.
 KW Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
 KW Transcription regulation.
 FT DOMAIN 580 670
 FT MOD_RES 705 705 Phosphotyrosine (by JAK) (By similarity).
 FT MOD_RES 727 727 Phosphoserine (By similarity).
 SQ SEQUENCE 770 AA; 88039 MW; D74A0C76954754ED CRC64;

Query Match 99.0%; Score 1374; DB 1; Length 770;
 Best Local Similarity 98.9%; Pred. No. 3e-77;
 Matches 268; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 RCLMBESRLQTAATAAQQGGGGAHPPTAAVTEKQOMLEOHLOVRRKRVODLEQMKVYE 60
 DB 107 RCLMBESRLQTAATAAQQGGGGAHPPTAAVTEKQOMLEOHLOVRRKRVODLEQMKVYE 166
 OY 61 NLQDDPFPNYKTLKSQGDMDLNGNNSVTRQKMQOLEOMLTALDQMRSTVSELAGLLS 120
 DB 167 NLQDDPFPNYKTLKSQGDMDLNGNNSVTRQKMQOLEOMLTALDQMRSTVSELAGLLS 226
 OY 121 AMEYVQKTLTDEELADWKRPELIACIGPPNICDRLNNWTTSLAESOLQTRQOIKKLEE 180
 DB 227 AMEYVQKTLTDEELADWKRPELIACIGPPNICDRLNNWTTSLAESOLQTRQOIKKLEE 286
 OY 181 LQOKVSYKGDPIVQHRPMLERIVLEFNNLKSAFVVERQPCMPHPRPIVITKGVQFT 240
 DB 287 LQOKVSYKGDPIVQHRPMLERIVLEFNNLKSAFVVERQPCMPHPRPIVITKGVQFT 346
 OY 241 TKVRLVKEPELANYQLKIKVCIDKDSGDVAA 271
 DB 347 TKVRLVKEPELANYQLKIKVCIDKDSGDVAA 377

RESULT 5

STAT3 BOVIN
 ID STAT3 BOVIN STANDARD; PRT; 770 AA.
 AC P61635;
 DT 05-JUN-2004 (Rel. 44, Created)
 DT 05-JUN-2004 (Rel. 44, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Signal transducer and activator of transcription 3.
 GN Name=STAT3;
 OS Bos taurus (Bovine).
 OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovinae; Bos.
 OX NCBI_TaxID=9913;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Mammary gland;
 RA Seyfert H.M., Wheeler T.T., Moolenaar A., Pitra C.;
 RT "The STAT3B-encoding gene was flipped across the STAT3/STAT5A-locus
 RT during ruminant evolution.";
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Transcription factor that binds to the interleukin-6
 CC (IL-6)-responsive elements identified in the promoters of various
 CC acute-phase protein genes (By similarity).
 CC -1- PATHWAY: Involved in the gp130-mediated signaling pathway.
 CC -1- SUBUNIT: Forms a homodimer or a heterodimer with a related family
 CC member (at least STAT1) (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
 CC in response to phosphorylation (By similarity).
 CC -1- SIMILARITY: Belongs to the transcription factor STAT family.
 CC -1- SIMILARITY: Contains 1 SH2 domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AJ620655; CAF06182.1; -.
 DR PROSITE: PS00001; SH2; 1.
 KW Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
 KW Transcription regulation.
 FT DOMAIN 580 670
 FT MOD_RES 705 705 Phosphotyrosine (by JAK) (By similarity).
 FT MOD_RES 727 727 Phosphoserine (By similarity).
 SQ SEQUENCE 770 AA; 87974 MW; 9CEB147C73B83274 CRC64;

Query Match 98.7%; Score 1370; DB 1; Length 770;
 Best Local Similarity 98.9%; Pred. No. 5.3e-77;
 Matches 268; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 RCLMBESRLQTAATAAQQGGGGAHPPTAAVTEKQOMLEOHLOVRRKRVODLEQMKVYE 60
 DB 107 RCLMBESRLQTAATAAQQGGGGAHPPTAAVTEKQOMLEOHLOVRRKRVODLEQMKVYE 166
 OY 61 NLQDDPFPNYKTLKSQGDMDLNGNNSVTRQKMQOLEOMLTALDQMRSTVSELAGLLS 120
 DB 167 NLQDDPFPNYKTLKSQGDMDLNGNNSVTRQKMQOLEOMLTALDQMRSTVSELAGLLS 226
 OY 121 AMEYVQKTLTDEELADWKRPELIACIGPPNICDRLNNWTTSLAESOLQTRQOIKKLEE 180
 DB 227 AMEYVQKTLTDEELADWKRPELIACIGPPNICDRLNNWTTSLAESOLQTRQOIKKLEE 286
 OY 181 LQOKVSYKGDPIVQHRPMLERIVLEFNNLKSAFVVERQPCMPHPRPIVITKGVQFT 240
 DB 287 LQOKVSYKGDPIVQHRPMLERIVLEFNNLKSAFVVERQPCMPHPRPIVITKGVQFT 346
 OY 241 TKVRLVKEPELANYQLKIKVCIDKDSGDVAA 271
 DB 347 TKVRLVKEPELANYQLKIKVCIDKDSGDVAA 377

```

RESULT 6
Q6DV79          PRELIMINARY;      PRT;      771 AA.
AC Q6DV79;
DT 01-OCT-2004 (TREMBlrel. 28, Created)
DT 01-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Signal transducer and activator of transcription 3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhou G.Y., Leung F.C.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY641397; AAT64887.1; -
SQ SEQUENCE 771 AA; 88174 MW; 71AC855C5DEC03E2 CRC64;

Query Match          97.4%; Score 1352; DB 2; Length 771;
Best Local Similarity 97.4%; Pred. No. 7e-76;
Matches 264; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 RCLMEESRLQTAATTAAGGGAHPATAVTEKQOMLEQHIQDVAKKRVQDLEQKKKVE 60
DB 107 RCLMEESRLQTAATTAAGGGAHTPTAAVTEKQOMLEQHIQDVAKKRVQDLEQKKKVE 166
QY 61 NLQDDPDPNYKTLKSGQMDLNGNNSVTRQKMOQLQMLTALDQMRISYSEIAGLLS 120
DB 167 NLQDDPDPNYKTLKSGQMDLNGNNSVTRQKMOQLQMLTALDQMRISYSEIAGLLS 226
QY 121 AMEYVQKTLTDEBLADWKRREPIACIGPPNICLDRLENWITSLAESQIQKQIKKEE 180
DB 227 AMEYVQKTLTDEBLADWKRREPIACIGPPNICLDRLENWITSLAESQIQKQIKKEE 286
QY 181 LQOKSYKGDPIVQHRPMLSEIRIVELFRNLMSAFVVERQPCMPHPDRPLVTKGVQFT 240
DB 287 LQOKSYKGDPIVQHRPMLSEIRIVELFRNLMSAFVVERQPCMPHPDRPLVTKGVQFT 346
QY 241 TKVRLVKEPELNYQIKIKVCIDKXSGDYAA 271
DB 347 TKVRLVKEPELNYQIKIKVCIDKXSGDYAA 377

RESULT 7
Q9PVX8          PRELIMINARY;      PRT;      769 AA.
AC Q9PVX8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Stat 3.
GN Name=stat 3;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20107399; PubMed=10642787;
RA Nishinakamura R., Matsunoto Y., Matsuda T., Arizumi T., Heike T.,
RA Asahina M., Yokota T.;
RA "Activation of Stat3 by cytokine receptor gp130 ventralizes Xenopus
RT embryos independent of BMP-4."
RL Dev. Biol. 216:481-490(1999).
DR EMBL; AB017701; BAA6061.1; -
DR HSRF; P42227; IBI.
RA GO; GO:0005633; C:nucleus; IEA.
RA GO; GO:0004871; F:signal transducer activity; IEA.
RA GO; GO:0003700; F:transcription factor activity; IEA.
RA GO; GO:0007242; P:intracellular signaling cascade; IEA.

```

```

DR GO; GO:006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PR00017; SH2; 1.
DR Pfam; PR01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 769 AA; 87974 MW; 0905C03263303069 CRC64;

Query Match          93.9%; Score 1304; DB 2; Length 769;
Best Local Similarity 93.0%; Pred. No. 6.8e-73;
Matches 252; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1 RCLMEESRLQTAATTAAGGGAHPATAVTEKQOMLEQHIQDVAKKRVQDLEQKKKVE 60
DB 107 RCLMEESRLQTAATTAAGGGAHPATAVTEKQOMLEQHIQDVAKKRVQDLEQKKKVE 166
QY 61 NLQDDPDPNYKTLKSGQMDLNGNNSVTRQKMOQLQMLTALDQMRISYSEIAGLLS 120
DB 167 NLQDDPDPNYKTLKSGQMDLNGNNSVTRQKMOQLQMLTALDQMRISYSEIAGLLS 226
QY 121 AMEYVQKTLTDEBLADWKRREPIACIGPPNICLDRLENWITSLAESQIQKQIKKEE 180
DB 227 AMEYVQKTLTDEBLADWKRREPIACIGPPNICLDRLENWITSLAESQIQKQIKKEE 286
QY 181 LQOKSYKGDPIVQHRPMLSEIRIVELFRNLMSAFVVERQPCMPHPDRPLVTKGVQFT 240
DB 287 LQOKSYKGDPIVQHRPMLSEIRIVELFRNLMSAFVVERQPCMPHPDRPLVTKGVQFT 346
QY 241 TKVRLVKEPELNYQIKIKVCIDKXSGDYAA 271
DB 347 NKVRLVKEPELNYQIKIKVCIDKXSGDYAA 377

RESULT 8
Q7ZXK3          PRELIMINARY;      PRT;      766 AA.
AC Q7ZXK3;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Stat3-A protein.
GN Name=stat3-A;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Embryo;
RA MEDLINE=22388257; PubMed=12477932;
RA Strusberg R.V., Reingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

```


RT and mouse cDNA sequence:"
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative";
 RL Dev. Dyn. 225:384-391 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC044717; AAA44717.1; -
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0004871; F:signal transducer activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR008967; PS3_like_DNA_bnd.
 DR InterPro; IPR00980; SH2.
 DR InterPro; IPR01217; STAT.
 DR Pfam; PF00117; SH2; 1.
 DR Pfam; PF01017; STAT_alpha; 1.
 DR Pfam; PF02864; STAT_bird; 1.
 DR Pfam; PF02865; STAT_int; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS50001; SH2; 1.
 SQ SEQUENCE 766 AA; 87599 MW; 31018A3321CCEB9C CRC64;

Query Match 92.7%; Score 1266; DB 2; Length 766;
 Best Local Similarity 91.1%; Pred. No. 8.9e-72;
 Matches 247; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

QY 1 RCLMEESRLTQTATAAOCGQANHTAAVTEKQOMLEOHLOVDRKRVODLEQKMYVE 60
 DB 107 RCLMEESRLTQTATAAOCGQANHTAAVTEKQOMLEOHLOVDRKRVODLEQKMYVE 166
 QY 61 NLQDDFDNFKYKLSQGMODLNGNNSVTRQKMOLEQMLTALDQMRRSIVSLAGL 120
 DB 167 NLQDDFDNFKYKLSQGMODLNGNNSVTRQKMOLEQMLTALDQMRRSIVSLAGL 226
 QY 121 AMEVYQKLTDEBELADMKRRPEIACIGGPNICLDRLNWITSLAESQLOTRQOIKLEE 180
 DB 227 AMEVYQKLTDEBELADMKRRPEIACIGGPNICLDRLNWITSLAESQLOTRQOIKLEE 286
 QY 181 LQOKSVYKGDPIVQHRPMLERIVLFRNLKSAFVERQPCMPHDPRLVIKTVQFT 240
 DB 287 LQOKSVYKGDPIVQHRPMLERIVLFRNLKSAFVERQPCMPHDPRLVIKTVQFT 346
 QY 241 TKVRLVVKPELNYQKIKVICIDKSGDVAA 271
 DB 347 NKVRLVVKPELNYQKIKVICIDKSGDVAA 377
 RESULT 9
 Q7ZTS5 PRELIMINARY; PRT; 414 AA.
 AC Q7ZTS5; 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DR Stat3 protein.
 GN Name=Stat3;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Actinoptera; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Cyprinodontiformes; Cyprinodontiformes; Cyprinodontiformes;
 OC Cyprinidae; Danio.
 OK NCBI_taxid=7955;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Whole body;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Hellon E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalski U., Smalins D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Whole body;
 RA Strausberg R.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC045276; AAA45276.1; -
 DR ZFIN; ZDB-GENE-980526-66; stat3.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0004871; F:signal transducer activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0007145; P:signal transduction; IEA.
 DR InterPro; IPR008967; PS3_like_DNA_bnd.
 DR InterPro; IPR01217; STAT.
 DR Pfam; PF01017; STAT_alpha; 1.
 DR Pfam; PF02864; STAT_bird; 1.
 DR Pfam; PF02865; STAT_int; 1.
 SQ SEQUENCE 414 AA; 48253 MW; 0FFD1B509B7526BD CRC64;

Query Match 85.8%; Score 1191.5; DB 2; Length 414;
 Best Local Similarity 84.6%; Pred. No. 3.3e-66;
 Matches 231; Conservative 22; Mismatches 17; Indels 3; Gaps 3;

QY 1 RCLMEESRLTQTATAAOCGQANHTAAVTEKQOMLEOHLOVDRKRVODLEQKMYVE 60
 DB 107 RCLMEESRLTQTATAAOCGQANHTAAVTEKQOMLEOHLOVDRKRVODLEQKMYVE 165
 QY 61 NLQDDFDNFKYKLSQGMODLNGNNSVTRQKMOLEQMLTALDQMRRSIVSLAGL 118
 DB 166 NLQDDFDNFKYKLSQGMODLNGNNSVTRQKMOLEQMLTALDQMRRSIVSLAGL 225
 QY 119 LSAEYVQKLTDEBELADMKRRPEIACIGGPNICLDRLNWITSLAESQLOTRQOIKLEE 178
 DB 226 LSAEYVQKLTDEBELADMKRRPEIACIGGPNICLDRLNWITSLAESQLOTRQOIKLEE 285
 QY 179 BELQOKSVYKGDPIVQHRPMLERIVLFRNLKSAFVERQPCMPHDPRLVIKTVQ 238
 DB 286 BELQOKSVYKGDPIVQHRPMLERIVLFRNLKSAFVERQPCMPHDPRLVIKTVQ 345
 QY 239 FTTKRLVVKPELNYQKIKVICIDKSGDVAA 271
 DB 346 FTTKRLVVKPELNYQKIKVICIDKSGDVAA 378
 RESULT 10
 Q6NY46 PRELIMINARY; PRT; 786 AA.
 AC Q6NY46; 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, last annotation update)
 DE Stat3 protein.
 GN Name=stat3;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullan S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
 Krzywinski M.I., Skalska U., Smalhus D.E., Scherch A., Schein J.E.,
 Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
 RA EMBL: BC068320; AAH68320.1;
 DR InterPro: IPR008967; P53_Like_DNA_bnd.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF01017; STAT alpha; 1.
 DR Pfam: PF02864; STAT bind; 1.
 DR Pfam: PF02865; STAT_int; 1.
 DR PROSITE: P55001; SH2; 1.
 SQ SEQUENCE 786 AA; 90039 MW; FC7371D0B0E5447E CRC64;
 Query Match 85.8%; Score 1191.5; DB 2; Length 786;
 Best Local Similarity 84.6%; Pred. No. 6.8e-66;
 Matches 231; Conservative 22; Mismatches 17; Indels 3; Gaps 3;
 QY 1 RCLMEBSRLLOTPAATPAAGGCGQANHPAAVTEKQOMLEOHLOVYKRVODLEQKKKVE 60
 DB 107 RCLMEBSRLLOTPAATPAAGGCGQANHPAAVTEKQOMLEOHLOVYKRVODLEQKKKVE 165
 QY 61 NIQDPDFENYKTLKSGGDM-ODLNGNNO-SVTRKKQOOLEOMITLADQMRSTVSISLAG 118
 DB 166 NIQDPDFENYKTLKSGGDM-ODLNGNNO-SVTRKKQOOLEOMITLADQMRSTVSISLAG 225
 QY 119 ISAMEVVOXTLDEBLADWKRREPEIACIGGPNPICDRLENNITSLAESQLOTRQOIKTL 178
 DB 226 ISAMPFVQGNLDEBLADWKRREPEIACIGGPNPICDRLENNITSLAESQLOTRQOIKTL 285
 QY 179 BELQOKVSYKGPPIVOHRPMEERIVELFRNLKSAFVVEROPCPMHEDRPLVITKGVQ 238
 DB 286 BELQOKVSYKGPPIVOHRPMEERIVELFRNLKSAFVVEROPCPMHEDRPLVITKGVQ 345
 QY 239 FTTKVRLLVKEPELNVOLKIKVICIDKSGDVAA 271
 DB 346 FTTKVRLLVKEPELNVOLKIKVICIDKSGDVAA 378

RESULT 11
 ID AAH68320 PRELIMINARY; PRT; 786 AA.
 AC AAH68320;
 DT 10-MAY-2004 (TrEMBLrel. 27, Created)
 DT 10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
 DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
 DT 10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
 DE Stat3 protein.
 GN Name=stat3;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullan S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
 Krzywinski M.I., Skalska U., Smalhus D.E., Scherch A., Schein J.E.,
 Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
 RA EMBL: BC068320; AAH68320.1;
 DR InterPro: IPR008967; P53_Like_DNA_bnd.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF01017; STAT alpha; 1.
 DR Pfam: PF02864; STAT bind; 1.
 DR Pfam: PF02865; STAT_int; 1.
 DR PROSITE: P55001; SH2; 1.
 SQ SEQUENCE 786 AA; 90039 MW; FC7371D0B0E5447E CRC64;
 Query Match 85.8%; Score 1191.5; DB 2; Length 786;
 Best Local Similarity 84.6%; Pred. No. 6.8e-66;
 Matches 231; Conservative 22; Mismatches 17; Indels 3; Gaps 3;
 QY 1 RCLMEBSRLLOTPAATPAAGGCGQANHPAAVTEKQOMLEOHLOVYKRVODLEQKKKVE 60
 DB 107 RCLMEBSRLLOTPAATPAAGGCGQANHPAAVTEKQOMLEOHLOVYKRVODLEQKKKVE 165
 QY 61 NIQDPDFENYKTLKSGGDM-ODLNGNNO-SVTRKKQOOLEOMITLADQMRSTVSISLAG 118
 DB 166 NIQDPDFENYKTLKSGGDM-ODLNGNNO-SVTRKKQOOLEOMITLADQMRSTVSISLAG 225
 QY 119 ISAMEVVOXTLDEBLADWKRREPEIACIGGPNPICDRLENNITSLAESQLOTRQOIKTL 178
 DB 226 ISAMPFVQGNLDEBLADWKRREPEIACIGGPNPICDRLENNITSLAESQLOTRQOIKTL 285
 QY 179 BELQOKVSYKGPPIVOHRPMEERIVELFRNLKSAFVVEROPCPMHEDRPLVITKGVQ 238
 DB 286 BELQOKVSYKGPPIVOHRPMEERIVELFRNLKSAFVVEROPCPMHEDRPLVITKGVQ 345
 QY 239 FTTKVRLLVKEPELNVOLKIKVICIDKSGDVAA 271
 DB 346 FTTKVRLLVKEPELNVOLKIKVICIDKSGDVAA 378

RESULT 12
 093599

ID 093599 PRELIMINARY; PRT; 806 AA.
AC 093599;
DT 01-NOV-1998 (Tremblrel. 03, Created)
DT 01-NOV-1998 (Tremblrel. 03, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Transcription factor.
GN Name=etac3;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RX [1]
RP SEQUENCE FROM N.A.
RA Oates A.C.;
RL Thesis (1998), University of Melbourne, Australia.
DR EMBL; AJ005693; CAA06677.1; -.
DR HSSP; P42227; 1801.
DR ZFIN; ZDB-GENE-980526-68; stat3.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009807; PS3_like_DNA_bnd.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 2.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 806 AA; 92151 MW; 74BC4EA401C3C942 CRC64;

Query Match 85.8%; Score 1191.5; DB 2; Length 806;
Best Local Similarity 84.6%; Pred. No. 7e-66;
Matches 231; Conservative 22; Mismatches 17; Indels 3; Gaps 3;

QY 1 RCLWESRLIQTAAATAAQQGGQANHPPTAAVTEKQMLEQHLQVRRKVDLEQMKVVE 60
DB 107 RCLWESRLIQT-ATTSODGSAHPPTGVTEKQMLEHNLQDIRRQVDEQMKMLE 165
QY 61 NLQDDPFPNFKTLKSGQDM-ODLNGNNO-SVTRQKMOQLBOMLTALDQMRISYSELGL 118
DB 166 NLQDDPFPNFKTLKSGQDM-ODLNGNNO-SVTRQKMOQLBOMLTALDQMRISYSELGL 225
QY 119 LSAMEYQKTLTDELDADWKRPRPIACTIGGPPNICDLRLNWITSLAESQLOTRQOIKKL 178
DB 226 LSAMDFOKNTLDELADWKRROQIACIGGPPNICDLRLNWITSLAESQLOTRQOIKKL 285
QY 179 BELQOKSYKGDPIVQHRPMLERIVELFRNLKSAFVERQPCPMHDPRLVIKTVQ 238
DB 286 BELQOKSYKGDPIVQHRPMLERIVELFRNLKSAFVERQPCPMHDPRLVIKTVQ 345
QY 239 FTTKVRLLVKEPELNYQKIKVCIDKSGDVAA 271
DB 346 FTTKVRLLVKEPELNYQKIKVCIDKSGDVAA 378

RESULT 13

Q6DVFP3 PRELIMINARY; PRT; 765 AA.

ID 06DVFP3
AC 06DVFP3;
DT 01-OCT-2004 (Tremblrel. 28, Created)
DT 01-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 01-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Signal transducer and activation of transcription factor 3.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu R., Hong Y.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF639947; AAT64912.1; -.
DR EMBL; AF639947; AAT64912.1; -.
SQ SEQUENCE 765 AA; 87566 MW; F5D01A08748BC703 CRC64;

Query Match 85.2%; Score 1182.5; DB 2; Length 765;
Best Local Similarity 83.9%; Pred. No. 2.4e-65;
Matches 229; Conservative 22; Mismatches 19; Indels 3; Gaps 3;

QY 1 RCLWESRLIQTAAATAAQQGGQANHPPTAAVTEKQMLEQHLQVRRKVDLEQMKVVE 60
DB 107 RCLWESRLIQT-ATTSODGSAHPPTGVTEKQMLEHNLQDIRRQVDEQMKMLE 165
QY 61 NLQDDPFPNFKTLKSGQDM-ODLNGNNO-SVTRQKMOQLBOMLTALDQMRISYSELGL 118
DB 166 NLQDDPFPNFKTLKSGQDM-ODLNGNNO-SVTRQKMOQLBOMLTALDQMRISYSELGL 225
QY 119 LSAMEYQKTLTDELDADWKRPRPIACTIGGPPNICDLRLNWITSLAESQLOTRQOIKKL 178
DB 226 LTAMDYQKNTLDELADWKRROQIACIGGPPNICDLRLNWITSLAESQLOTRQOIKKL 285
QY 179 BELQOKSYKGDPIVQHRPMLERIVELFRNLKSAFVERQPCPMHDPRLVIKTVQ 238
DB 286 BELQOKSYKGDPIVQHRPMLERIVELFRNLKSAFVERQPCPMHDPRLVIKTVQ 345
QY 239 FTTKVRLLVKEPELNYQKIKVCIDKSGDVAA 271
DB 346 FTTKVRLLVKEPELNYQKIKVCIDKSGDVAA 378

RESULT 14

Q6GUE7 PRELIMINARY; PRT; 785 AA.

ID 06GUE7
AC 06GUE7;
DT 05-JUN-2004 (Tremblrel. 27, Created)
DT 05-JUN-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUN-2004 (Tremblrel. 27, Last annotation update)
DE Signal transducer and activator of transcription 3 isoform 1.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OC NCBI_TaxID=8090;
RX [1]
RP SEQUENCE FROM N.A.
RA Liu R., Hong Y.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY641434; AAT6364.1; -.
DR InterPro; IPR008967; PS3_like_DNA_bnd.
DR InterPro; IPR009807; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 785 AA; 89643 MW; 81FC21BDE27DE938 CRC64;

Query Match 85.2%; Score 1182.5; DB 2; Length 785;
Best Local Similarity 83.9%; Pred. No. 2.5e-65;
Matches 229; Conservative 22; Mismatches 19; Indels 3; Gaps 3;

QY 1 RCLWESRLIQTAAATAAQQGGQANHPPTAAVTEKQMLEQHLQVRRKVDLEQMKVVE 60
DB 107 RCLWESRLIQT-ATTSODGSAHPPTGVTEKQMLEHNLQDIRRQVDEQMKMLE 165
QY 61 NLQDDPFPNFKTLKSGQDM-ODLNGNNO-SVTRQKMOQLBOMLTALDQMRISYSELGL 118
DB 166 NLQDDPFPNFKTLKSGQDM-ODLNGNNO-SVTRQKMOQLBOMLTALDQMRISYSELGL 225

```

QY 119 LSAMEYVQKTLTDEEADWKREPELACIGPPNICDRLIEMWITSIAESQLOTRQIKKL 178
DB 226 LTAMDYVQKNTLDEEADWKRRQOQIACIGPPNICDRLIETWITSIAESQLOTRQIKKL 285
QY 179 ERIQOKVSYKGPPIVQHRPMLERIVELFRNLKMSAFVVEROPCMHPDRPLVITKGVQ 238
DB 286 ERIQOKVSYKGPPIVQHRPMLERIVELFRNLKMSAFVVEROPCMHPDRPLVITKGVQ 345
QY 239 FTTKVRLLVKPELNVQKIKVCIDKDSGDVAA 271
DB 346 FTKKVRLLVKPELNVQKIKVCIDKDSGDVAA 378

```

RESULT 15

```

Q90Y16 PRELIMINARY; FRT; 764 AA.
AC Q90Y16;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE STAT3.
OS Tetraodon fluviatilis (Puffer fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
NCBI_Taxid=47145;
RN [1]
RP SEQUENCE FROM N.A.
RA Sung S.C., Fan T.J., Liu J.H., Chou C.M., Huang C.J.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
EMBL: AF307106; AAL09415.1; -
DR HSSP; P42227; 1BG1.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008967; P53_like_CNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR01217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
SQ SEQUENCE 764 AA; 87411 MW; E661FE18BEPDBE CRC64;

```

```

Query Match 84.0%; Score 1165.5; DB 2; Length 764;
Best Local Similarity 83.2%; Pred. No. 2.7e-64;
Matches 227; Conservative 24; Mismatches 19; Indels 3; Gaps 3;

```

```

QY 1 RCIMESRLLQTAATTAACGGQANHPATAVTEKQOMLEQHLQDYKRYVQDLEQKKYVE 60
DB 107 RCIMESRLLQTAATTAACGGQANHPATAVTEKQOMLEQHLQDYKRYVQDLEQKKYVE 165
QY 61 NIQDDFDENVYKTLKSGDM-QDINGNQ-SVTRQKQOLEOMLTALDQWRSIVSELGL 118
DB 166 NIQDDFDENVYKTLKSGDM-QDINGNQ-SVTRQKQOLEOMLTALDQWRSIVSELGL 225
QY 119 LSAMEYVQKTLTDEEADWKRRPELACIGPPNICDRLIEMWITSIAESQLOTRQIKKL 178
DB 226 LTAMDYVQKNTLDEEADWKRRQOQIACIGPPNICDRLIETWITSIAESQLOTRQIKKL 285
QY 179 ERIQOKVSYKGPPIVQHRPMLERIVELFRNLKMSAFVVEROPCMHPDRPLVITKGVQ 238
DB 286 ERIQOKVSYKGPPIVQHRPMLERIVELFRNLKMSAFVVEROPCMHPDRPLVITKGVQ 345
QY 239 FTTKVRLLVKPELNVQKIKVCIDKDSGDVAA 271
DB 346 FTKKVRLLVKPELNVQKIKVCIDKDSGDVAA 378

```

Search completed: December 2, 2004, 23:52:25
 Job time : 196 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2004, 23:38:03; Search time 39 Seconds

(without alignments)
668.583 Million cell updates/sec

Title: US-10-090-185-9

Perfect score: 1388

Sequence: 1 RCLWESRLIQTATAAQAQG.....LNYQKIKVICIDKDSGVAA 271

Scoring table: BLOSUM62

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1377	99.2	770	2	IGSF3 p91-related
2	1372	98.8	770	2	DNA-binding protei
3	601.5	43.3	739	2	interferon-depende
4	582.5	42.0	748	2	gamma-interferon a
5	364	26.2	851	2	interferon alpha-1
6	288.5	20.8	786	2	mammary gland fact
7	284.5	20.5	793	2	mammary gland fact
8	277.5	20.0	794	2	transcription acti
9	252	18.2	794	2	mammary gland fact
10	129	9.3	837	2	DNA-binding Protei
11	129	9.3	837	2	interleukin-4-indu
12	128	9.2	1208	2	chromosome segrega
13	126	9.1	978	2	conserved hypotet
14	125.5	9.0	1166	2	hypothetical prote
15	125	9.0	464	2	microtubule bindin
16	124	8.9	533	2	hypothetical prote
17	122.5	8.8	857	2	median body protei
18	122	8.8	1972	1	myosin heavy chain
19	120	8.6	734	2	hypothetical prote
20	120	8.6	1509	1	myosin heavy chain
21	118.5	8.5	924	2	myosin heavy chain
22	118.5	8.5	2094	2	myosin heavy chain
23	118	8.5	284	2	myosin heavy chain
24	118	8.5	2253	2	myosin heavy chain
25	118	8.5	2442	2	myosin heavy chain
26	117.5	8.5	2007	1	myosin heavy chain
27	117	8.4	1818	1	myosin heavy chain
28	117	8.4	1938	1	myosin heavy chain
29	117	8.4	1972	2	myosin heavy chain

30	116.5	8.4	1300	2	153799	CG1 protein - huma
31	116.5	8.4	1356	2	S32763	kinectin 1 - huma
32	116	8.4	289	2	S51193	epimorphin - mouse
33	116	8.4	946	2	S28061	SCP1 protein - rat
34	116	8.4	986	2	T10754	cis-Golgi matrix p
35	115.5	8.3	1178	2	S30431	MSP-300 protein -
36	115	8.3	1164	2	T24806	hypothetical prote
37	114.5	8.2	1690	2	T13030	microtubule bindin
38	114.5	8.2	1999	1	S21801	myosin heavy chain
39	114.5	8.2	3187	2	JC5837	364k Golgi complex
40	114	8.2	1008	2	AE2304	hypothetical prote
41	114	8.2	1410	1	A57013	early endosome ant
42	114	8.2	1961	1	A61231	myosin heavy chain
43	114	8.2	1963	1	MMKW	myosin heavy chain
44	114	8.2	1976	2	A59252	myosin heavy chain
45	114	8.2	2663	1	S28261	centromere protein

ALIGNMENTS

RESULT 1

IGSF3 p91-related transcription factor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004

C:Accession: I49508; I49009

R:AKira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Su

Cell 77, 63-71, 1994

A:Title: Molecular cloning of APRF, a novel IFN- γ -stimulated gene factor 3 p91-related tr

A:Reference number: A54444; MUID:94208062; PMID:7512451

A:Accession: I49508

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-770 <RES>

A:Cross-references: UNIPROT:P42227; GB:I29278; NID:G476715; PID:AAA37254.1; PID:G47671

R:Raz, R.; Durbin, J.B.; Levy, D.B.

J. Biol. Chem. 269, 24391-24395, 1994

A:Title: Acute phase response factor and additional members of the interferon-stimulate

A:Reference number: I49009; MUID:95014185; PMID:7523373

A:Accession: I49009

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-393, 'M', 395-700, 702-770 <RES>

A:Cross-references: EMBL:U08378; NID:G473889; PID:AAA56668.1; PID:G473890

C:Genetics:

A:Gene: APRF

C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 99.2%; Score 1377; DB 2; Length 770;

Best Local Similarity 99.3%; Pred. No. 4.8e-81;

Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY	1	RCLWESRLIQTATAAQAQGQANHPPRAVTEQQMLFQHLQVRRVQDLEQMKYVE	60
DB	107	RCLWESRLIQTATAAQAQGQANHPPRAVTEQQMLFQHLQVRRVQDLEQMKYVE	166
QY	61	NLQDDPFPNYKTLSSQGMODLNGNOSVTRQKQQLQEQMLTALDQMRSLVSELAGLLS	120
DB	167	NLQDDPFPNYKTLSSQGMODLNGNOSVTRQKQQLQEQMLTALDQMRSLVSELAGLLS	226
QY	121	AMEVYQKTLDBELADWKRREPIACIGGPPNICDRLQEMWITSLAESQLQTRQOIKLEE	180
DB	227	AMEVYQKTLDBELADWKRREPIACIGGPPNICDRLQEMWITSLAESQLQTRQOIKLEE	286
QY	181	LQQRVSYKGDPIVQHRPMLERIVLFPNLKMSAFVVRQCMRPHRPRPLVTKGVQFT	240
DB	287	LQQRVSYKGDPIVQHRPMLERIVLFPNLKMSAFVVRQCMRPHRPRPLVTKGVQFT	346
QY	241	TKVRLLVFPELANTQKIKVICIDKDSGVAA 271	
DB	347	TKVRLLVFPELANTQKIKVICIDKDSGVAA 377	

A:Accession: A46160
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA; protein
 A:Residues: 1-851 <FUI>
 A:Cross-references: UNIPROT:P52630
 A:Note: sequence extracted from NCBI backbone (NCBI:P110820)
 R:Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell, J.E.
 submitted to the EMBL Data Library, December 1994
 A:Reference number: S71908
 A:Accession: S71908
 A:Molecule type: DNA
 A:Residues: 1-851 <YAN>
 A:Cross-references: EMBL:U18671; NID:G1293919; P1DN:AAA98760.1; PID:G1293920
 R:Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell Jr., J.E.
 Nucleic Acids Res. 23, 459-463, 1995
 A:Title: The genomic structure of the STAT genes: multiple exons in coincident sites in
 A:Reference number: S53873; MUID:95192056; PMID:7885841
 A:Accession: S53873
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-196;392-591;684-730 <YAN>
 A:Cross-references: EMBL:U18671
 C:Genetics:
 A:Gene: stat2
 A:introns: 44/2; 95/3; 127/3; 157/3; 183/1; 211/3; 261/2; 314/2; 345/2; 365/2; 372/2; 40
 C:Superfamily: human signal transducer and transcription activator STAT5A
 C:Keywords: signal transduction; transcription regulation

Query Match 26.2%; Score 364; DB 2; Length 851;
 Best Local Similarity 33.7%; Pred. No. 6.5e-16;
 Matches 89; Conservative 58; Mismatches 107; Indels 10; Gaps 6;

QY 3 LMESRLIQTAATAAQQGQANHPAAVTEKQ-MLEOHLDVYKRVODLEQKAVEN 61
 DB 112 LLEERLIIQAPRAQLEQGE--PVLTERPVESQQHEISRLIDLAEMKLVKISQKLD 168
 QY 62 LQDDPDPYKTLKSGQMDLNGNOSTRQKMOLEMTALDQMRISYSELGLISA 121
 DB 169 QQDVCFPRYK-IOAKGKTPSIDPH--QTKRQKI--LQETINELDRKREVIDASKALIGR 223
 QY 122 MEVQKTLTDELDMDKRRPFIACIGPPNICLDRLEWINSLSAQQTQKKEEL 181
 DB 224 LTTLEILL-PLKEWKAQQAQACIRAPIDHGLELFTWFTAGALLFHLRLKLEKGL 282
 QY 182 QQKSYKGDPIVQHRPMLEERIVLEFRNIMKSAFVPERQCPMPHDPRLIKTGVOFTT 241
 DB 283 SCIVSYQDDPLTKGVDLNAQVTELLQRLHRAFAVFEVQPCMPQPHRPLIKTGSKRTV 342
 QY 242 KVRLLVKEPELANTQIKIVCTDKD 265
 DB 343 RTRLLVRLQEGNESILVVEISIDRN 366

RESULT 6
 149274
 mammary gland factor - mouse
 N:Alternate names: STAT5 protein homolog p80
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jul-1996 #sequence revision 02-Jul-1996 #text_change 09-Jul-2004
 C:Accession: 149274; S54773; S54727
 R:Lin, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hemmighausen, L.
 Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
 A:Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved in
 A:Reference number: 149273; MUID:96004632; PMID:7568026
 A:Accession: 149274
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-786 <RBS>
 A:Cross-references: UNIPROT:P42232; UNIPROT:Q9JMK1; EMBL:U21110; NID:G747973; P1DN:AA052
 R:Mul, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
 EMBO J. 14, 1166-1175, 1995
 A:Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin
 A:Reference number: S54772; MUID:95237198; PMID:7720707

A:Accession: S54773
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-432, 'E', 434-786 <MUI>
 A:Cross-references: EMBL:Z48539; NID:G758635; P1DN:CA08420.1; PID:G758636
 R:Azam, M.; Erdjument-Bromage, H.; Kreider, B.L.; Xia, M.; Queller, F.; Basu, R.; Saris,
 EMBO J. 14, 1402-1411, 1995
 A:Title: Interleukin-3 signals through multiple isoforms of Stat5.
 A:Reference number: S54725; MUID:95246733; PMID:7537213
 A:Accession: S54727
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-432, 'E', 434-786 <AZA>
 C:Genetics:
 A:Gene: Stat5b
 C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 20.8%; Score 288.5; DB 2; Length 786;
 Best Local Similarity 30.0%; Pred. No. 4.2e-11;
 Matches 79; Conservative 46; Mismatches 109; Indels 29; Gaps 7;

QY 1 RC---LMESRLIQTAATAAQQGQANHPAAVTEKQMLEOHLDVYKRVODLEQK 56
 DB 107 RCIRHILYNEQRLVREANNSSPAGS---LADAMSQKLDQINQTFEELRLITDTENEL 162
 QY 57 KVENLQDDPDPFNY-KTLKSGQMDLNGN-----QSVTRQKMOLEQML---TAL 104
 DB 163 KKLQQTQEVFIIOQESLRIDQAPRAQLEQLNPOERMSRFTLQKQVSLERLWQREAGTL 222
 QY 105 DQMRISYSELGLISAMEYQKTLTDELDMDKRRPFIACIGPPNICLDRLEWITS 164
 DB 223 QQVEVELAEKHQKTLQRLRKQOTIILDELLQWRKROQLAANGPPEGSIDLVLQGCERKL 282
 QY 165 AESQLQTRQOIKKEELQKSYKGDPIVQHRPMLEERIVLEFRNIMKSAFVPERQCPMP 224
 DB 283 ARIITWQNEQQRREHLLQOQPIPG-PVEEMLAEVNATITDIALVSTFITEKOP--- 338
 QY 225 MHPDRPLVKTGVQFTTKVRLV 247
 DB 339 -----PQVLKQTQTKFAATVRLV 356

RESULT 7
 S54772
 mammary gland factor - mouse
 N:Alternate names: stat5 protein
 C:Species: Mus musculus (house mouse)
 C:Date: 27-Oct-1995 #sequence revision 03-Nov-1995 #text_change 09-Jul-2004
 C:Accession: S54772; S54773
 R:Mul, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
 EMBO J. 14, 1166-1175, 1995
 A:Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin
 A:Reference number: S54772; MUID:95237198; PMID:7720707
 A:Accession: S54772
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-793 <MUI>
 A:Cross-references: UNIPROT:P42230; UNIPROT:Q9JIA0; EMBL:Z48538; NID:G758633; P1DN:CA08
 R:Lin, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hemmighausen, L.
 Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
 A:Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved
 A:Reference number: 149273; MUID:96004632; PMID:7568026
 A:Accession: 149273
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-793 <RBS>
 A:Cross-references: EMBL:U21103; NID:G747971; P1DN:AAA0590.1; PID:G747972
 C:Genetics:
 A:Gene: Stat5a
 C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 20.5%; Score 284.5; DB 2; Length 793;
 Best Local Similarity 29.7%; Pred. No. 7.7e-11;

